

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:49:28 ; Search time 43.8146 Seconds
(without alignments)
9625.515 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236
Perfect score: 51
Sequence: 1 acagaactggcaaggca.....agggcctgtcaccacagga 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	365	3	US-09-925-299-668 Sequence 668, App
2	51	100.0	365	3	US-09-925-299-668 Sequence 668, App
3	51	100.0	1777	8	US-10-623-108-7 Sequence 7, Appli
4	51	100.0	1837	8	US-10-623-108-5 Sequence 5, Appli
5	50	98.0	13808	6	US-09-764-877-3351 Sequence 3351, Ap
6	50	98.0	13808	6	US-10-242-515-3351 Sequence 3351, Ap
7	25.6	50.2	9916	3	US-09-816-095-3 Sequence 3, Appli
8	25.6	50.2	9916	3	US-10-634-905-3 Sequence 275780,
9	24.8	48.6	506	5	US-10-027-632-275790 Sequence 275790,
10	24.8	48.6	506	6	US-10-027-632-275790 Sequence 650600,
11	24.8	48.6	594	4	US-09-925-065A-650600 Sequence 21955, A
12	24.8	48.6	992	8	US-10-357-930-21955 Sequence 27809, A
13	24.8	48.6	992	8	US-10-357-930-27809 Sequence 3069, Ap
14	24.4	47.8	331	3	US-09-764-877-3070 Sequence 3070, Ap
15	24.4	47.8	331	3	US-10-242-515-3069 Sequence 3070, Ap
16	24.4	47.8	331	6	US-10-242-515-3070 Sequence 1105, Ap
17	24.4	47.1	582	8	US-10-774-355A-1105 Sequence 97357, A
18	24	47.1	2112	5	US-10-027-632-97357 Sequence 97357, A
19	24	47.1	2112	5	US-10-027-632-97358 Sequence 97358, A
20	24	47.1	2112	5	US-10-027-632-97359 Sequence 97359, A
21	24	47.1	2112	5	US-10-027-632-97359 Sequence 103733,
22	24	47.1	2112	5	US-10-027-632-103733 Sequence 103734,
23	24	47.1	2112	5	US-10-027-632-103734

C 24	24	47.1	2112	5	US-10-027-632-103735	Sequence 103735,
C 25	24	47.1	2112	6	US-10-027-632-97357	Sequence 97357, A
C 26	24	47.1	2112	6	US-10-027-632-97358	Sequence 97358, A
C 27	24	47.1	2112	6	US-10-027-632-97359	Sequence 97359, A
C 28	24	47.1	2112	6	US-10-027-632-103733	Sequence 103733,
C 29	24	47.1	2112	6	US-10-027-632-103734	Sequence 103734,
C 30	24	47.1	2112	6	US-10-027-632-103735	Sequence 103735,
C 31	24	47.1	51657	6	US-10-057-475B-10475	Sequence 10475, A
C 32	24	47.1	51657	6	US-10-154-884B-10475	Sequence 10475, A
C 33	23.8	46.7	346	8	US-10-425-115-173660	Sequence 173660,
C 34	23.8	46.7	1312	5	US-10-106-698-2115	Sequence 2115, Ap
C 35	23.8	46.7	1562	5	US-10-027-632-255989	Sequence 255989,
C 36	23.8	46.7	1562	6	US-10-027-632-255989	Sequence 255989,
C 37	23.8	46.7	3863	5	US-10-103-313-132	Sequence 132, App
C 38	23.8	46.7	247862	7	US-10-235-192A-28	Sequence 28, Appli
C 39	23.6	46.3	303	6	US-10-006-285-5	Sequence 5, Appli
C 40	23.6	46.3	506	6	US-10-029-386-6668	Sequence 6668, Ap
C 41	23.6	46.3	7862	8	US-10-723-860-4981	Sequence 4981, Ap
C 42	23.6	46.3	64011	8	US-10-719-993-6991	Sequence 6991, Ap
C 43	23.6	46.3	85571	8	US-10-719-993-6778	Sequence 6778, Ap
C 44	23.4	45.9	405	8	US-10-425-115-33612	Sequence 33612, A
C 45	23.4	45.9	17177	6	US-10-225-810-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-925-299-668
Sequence 668, Application US/09925299
Ratseq No. US2002005627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PAL02
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 668
LENGTH: 365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (172)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (239)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (243)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (244)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (329)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (330)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (353)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature

LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (362)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-299-668

Query Match 100.0%; Score 51; DB 3; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 32 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 82

RESULT 2

US-09-925-299-668
; Sequence 668, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 668
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (172)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (239)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (243)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (244)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (329)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (330)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (353)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (358)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (362)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-299-668

Query Match 100.0%; Score 51; DB 3; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 32 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 82

RESULT 3

US-10-623-108-7
; Sequence 7, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHWO
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-623-108-7

Query Match 100.0%; Score 51; DB 8; Length 1777;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 1126 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 1176

RESULT 4

US-10-623-108-5
; Sequence 5, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHWO
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-623-108-5

Query Match 100.0%; Score 51; DB 8; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 51
Db 1186 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCCTCTGTCAACCCAGGA 1236

RESULT 5

US-09-764-877-3351
; Sequence 3351, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3351

```
; LENGTH: 13808
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3351

Query Match      98.0%; Score 50; DB 3; Length 13808;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 50
    |||||
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 6424

RESULT 6
US-10-242-515-3351
; Sequence 3351, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3351
; LENGTH: 13808
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-3351

Query Match      98.0%; Score 50; DB 6; Length 13808;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 50
    |||||
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 6424

RESULT 7
US-09-816-095-3
; Sequence 3, Application US/09816095
; Patent No. US20020137164A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weinu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
```

```
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(99916)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match      50.2%; Score 25.6; DB 3; Length 99916;
Best Local Similarity 70.8%; Pred. No. 6.2;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAG 49
    |||||
Db 1819 CAGACCCGGGAGAGCCAAAGGGTCCCTCCGGGCTCCCTCTCGCAG 1866

RESULT 8
US-10-634-905-3
; Sequence 3, Application US/10634905
; Publication No. US20040067225A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weinu et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001147-DIV
; CURRENT APPLICATION NUMBER: US/10/634,905
; CURRENT FILING DATE: 2003-08-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(99916)
; OTHER INFORMATION: n = A,T,C or G
US-10-634-905-3

Query Match      50.2%; Score 25.6; DB 7; Length 99916;
Best Local Similarity 70.8%; Pred. No. 6.2;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAG 49
    |||||
Db 1819 CAGACCCGGGAGAGCCAAAGGGTCCCTCCGGGCTCCCTCTCGCAG 1866

RESULT 9
US-10-027-632-275790
; Sequence 275790, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
```

```

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275790
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-275790

Query Match      48.6%; Score 24.8; DB 5; Length 506;
Best Local Similarity 72.7%; Pred. No. 8;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      8  TGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGGA 51
Db      105 TGGCAAGGTGCCAGAGGTCCTTAATCGCATCTGTCTCCAGGA 148

RESULT 10
US-10-027-632-275790
; Sequence 275790, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275790
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-275790

Query Match      48.6%; Score 24.8; DB 6; Length 506;
Best Local Similarity 72.7%; Pred. No. 8;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      8  TGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGGA 51
Db      105 TGGCAAGGTGCCAGAGGTCCTTAATCGCATCTGTCTCCAGGA 148

RESULT 11
US-09-925-065A-650600
; Sequence 650600, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single

```

```

; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 650600
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-650600

Query Match      48.6%; Score 24.8; DB 4; Length 594;
Best Local Similarity 72.7%; Pred. No. 8.1;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      8  TGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCCACCAGGA 51
Db      85  TGGCAAGGTGCCAGAGGTCCTTAATCGCATCTGTCTCCAGGA 128

RESULT 12
US-10-357-930-21955
; Sequence 21955, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Ondege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21955
; LENGTH: 992
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 992
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-21955

Query Match      48.6%; Score 24.8; DB 8; Length 992;
Best Local Similarity 72.7%; Pred. No. 8.5;

```


Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCA 44
Db 624 ACTGCACTGTCAAGGCTGCAAGAGGCTCTGAGGGCTTCTGACA 667

RESULT 13

US-10-357-930-27809
; Sequence 27809, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27809
; LENGTH: 992
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 992
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-27809

Query Match 48.6%; Score 24.8; DB 8; Length 992;
Best Local Similarity 72.7%; Pred. No. 8.5;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCA 44
Db 624 ACTGCACTGTCAAGGCTGCAAGAGGCTCTGAGGGCTTCTGACA 667

RESULT 14

US-09-764-877-3069
; Sequence 3069, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3069
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-764-877-3069

Query Match 47.8%; Score 24.4; DB 3; Length 331;
Best Local Similarity 73.8%; Pred. No. 11;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTTC 43
Db 244 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 285

RESULT 15

US-09-764-877-3070
; Sequence 3070, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3070
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3070

Query Match 47.8%; Score 24.4; DB 3; Length 331;
Best Local Similarity 73.8%; Pred. No. 11;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTTC 43
Db 244 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 285

Search completed: February 7, 2006, 15:20:14
Job time : 44.8146 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:29:45 ; Search time 200.677 Seconds
(without alignments)
11890.431 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236
Perfect score: 51
Sequence: 1 acgaactggcaagagca.....aggcctctgtcaccaggga 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 2393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*
- 10: gb_est10.*
- 11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	555	2	BE383357 601298273
2	51	100.0	555	3	EM702058 UI-E-CQ1
3	51	100.0	555	6	CB156158 K-EST0214
4	51	100.0	653	2	BE891264 601432092
5	51	100.0	673	3	BM041252 603614446
6	51	100.0	673	6	CA446812 UI-H-ED1
7	51	100.0	773	3	BM041386 603614446
8	51	100.0	864	1	AL576621 AL576621
9	46.8	91.8	307	9	B50582 CIT-RSP-494
10	37.4	73.3	520	5	BQ640705 he32b09.Y
11	37.4	73.3	855	2	BF675069 602136667
12	28.4	55.7	500	1	AA324242 EST27116
13	27.8	54.5	412	7	CV358408 PMO-AN008
14	27	52.9	680	2	BE385462 601276049
15	26.6	52.2	698	3	B1869166 603395496
16	26.4	51.8	541	10	CW976199 A1AA-aab2
17	26	51.0	389	5	BY064864 BY064864
18	26	51.0	479	1	A1615963 mg61a04.Y
19	25.8	50.6	371	9	AQ207132 HS 3235.A
20	25.8	50.6	429	7	CV320055 CM2-UP008
21	25.8	50.6	536	9	AQ428895 C1BT1-E1
22	25.4	49.8	620	7	CN244360 EST010237

23	25.4	49.8	888	10	CNS010A0	AL153256 Anopheles
24	25	49.0	427	7	CK607655	CK607655 Controlb
25	25	49.0	535	8	CV999827	CV999827 iV53007.9
26	25	49.0	687	6	CA758516	CA758516 OF13B01-T
27	25	49.0	699	8	DN106444	DN106444 1103825.M
28	25	49.0	736	3	BP437662	BP437662 BP437662
29	25	49.0	804	2	BI182540	BI182540 UNL-P-FN-
30	25	49.0	856	7	CJ030128	CJ030128 CJ030128
31	24.8	48.6	428	6	CB298831	CB298831 220018.re
32	24.8	48.6	428	9	AQ877413	AQ877413 HS 2146.A
33	24.8	48.6	469	1	AI450334	AI450334 MU27e09.X
34	24.8	48.6	645	1	AL045331	AL045331 DKFZp434A
35	24.8	48.6	660	11	DE060758	DE060758 Oryzias.l
36	24.6	48.2	719	2	BI175907	BI175907 EST516990
37	24.6	48.2	764	3	BM112406	BM112406 EST559942
38	24.4	47.8	289	1	AW888774	AW888774 RCO-NT001
39	24.4	47.8	342	2	BF369448	BF369448 RCO-GN009
40	24.4	47.8	537	2	BG552612	BG552612 df44g10.x
41	24.4	47.8	562	11	TA269609Q	AL488207 T. brucei
42	24.4	47.8	697	9	CE237026	CE237026 tigr-g88-
43	24.4	47.8	813	5	BU373301	BU373301 603808454
44	24.4	47.8	836	5	BU453132	BU453132 603218119
45	24.4	47.8	869	5	BU456559	BU456559 603215358

ALIGNMENTS

RESULT 1	BE383357	601298273F1	NIH_MGC_19	Homo sapiens	cDNA clone IMAGE:3628314 5', mRNA linear EST 21-JUL-2000
LOCUS	BE383357	mRNA sequence.			
DEFINITION	BE383357	BE383357.1	GI:9328722		
ACCESSION	BE383357	EST.			
VERSION	BE383357.1	GI:9328722			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 555)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

ORIGIN

```
Query Match      100.0%; Score 51; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.6e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCTCTGTGTCACCCAGGA 51
    |||||||
Db 309 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCTCTGTGTCACCCAGGA 359

RESULT 2
LOCUS      BM702058                555 bp      mRNA      linear      EST 28-FEB-2002
DEFINITION UI-E-CQ1-ay-e-05-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone
VERSION     UI-E-CQ1-ay-e-05-0-UI 5', mRNA sequence.
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
REFERENCE   1 (bases 1 to 555)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
PUBMED     889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source           1..555
                     /lab_host="DHI0B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-E-CQ1"
                     /notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site 1: EcoR I; Site 2: Not I;
                     UI-E-CQ1 is a normalized cDNA library containing the
                     following tissue(s): optic nerve. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     CCATTAAGTG. This library was created for the program, Gene
                     Discovery in the Visual System, supported by National Eye
                     Institute (NEI)."
```

ORIGIN

```
Query Match      100.0%; Score 51; DB 3; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
```

```
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCTCTGTGTCACCCAGGA 51
    |||||||
Db 356 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCTCTGTGTCACCCAGGA 406

RESULT 3
LOCUS      CB156158                575 bp      mRNA      linear      EST 29-JAN-2003
DEFINITION K-EST0214813 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-2-C09
VERSION     CB156158.1 GI:28141276
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominiidae; Homo.
REFERENCE   1 (bases 1 to 575)
AUTHORS    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 2 row: C column: 09
            High quality sequence stop: 575.
FEATURES             Location/Qualifiers
     source           1..575
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L18POOL1n1-2-C09"
                     /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
                     /lab_host="Trop10P"
                     /clone_lib="L18POOL1n1"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
                     Site 2: NotI; The library was contributed by The Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
ORIGIN

Query Match      100.0%; Score 51; DB 6; Length 575;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCTCTGTGTCACCCAGGA 51
    |||||||
Db 222 ACAGAACTGGCAAGGCAAGAGGTCACTGAGGGCTCTGTGTCACCCAGGA 272

RESULT 4
LOCUS      BE891264                653 bp      mRNA      linear      EST 20-OCT-2000
DEFINITION 601432092F1 NIH MGC_72 Homo sapiens cDNA clone IMAGE:3917255 5',
VERSION     BE891264
KEYWORDS    mRNA sequence.
SOURCE      BE891264.1 GI:10350423
            EST.
            Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```

REFERENCE
AUTHORS      1 (bases 1 to 653)
TITLE        NIH-MGC http://mgs.nci.nih.gov/
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1AM9743 row: h column: 24
              High quality sequence stop: 602.

FEATURES
source
1. .653
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:3917255"
   /tissue_type="melanotic melanoma"
   /lab_host="DH10B (phage-resistant)"
   /clone_lib="NIH MGC 72"
   /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
   Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
   Average insert size 2 kb. Library constructed by Life
   Technologies."

ORIGIN
Query Match      100.0%; Score 51; DB 2; Length 653;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCCTCTGTCAACCAGGA 51
Db 416 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCCTCTGTCAACCAGGA 466

RESULT 5
BM041252
LOCUS      603161446F1 NIH_MGC_108 Homo sapiens cDNA clone IMAGE:5420314 5',
DEFINITION mRNA sequence.
ACCESSION  BM041252
VERSION     1 GI:16770519
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.

REFERENCE
1 (bases 1 to 673)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1874 row: 1 column: 11
High quality sequence stop: 648.

FEATURES
source
1. .673
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 51; DB 2; Length 653;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCCTCTGTCAACCAGGA 51
Db 416 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCCTCTGTCAACCAGGA 466

RESULT 5
BM041252
LOCUS      603161446F1 NIH_MGC_108 Homo sapiens cDNA clone IMAGE:5420314 5',
DEFINITION mRNA sequence.
ACCESSION  BM041252
VERSION     1 GI:16770519
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.

REFERENCE
1 (bases 1 to 673)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1874 row: 1 column: 11
High quality sequence stop: 648.

FEATURES
source
1. .673
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"

```

```

/clone="IMAGE:5420314"
/tissue_type="Wilms' tumor, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 108"
/note="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match      100.0%; Score 51; DB 3; Length 673;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCCTCTGTCAACCAGGA 51
Db 161 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCCTCTGTCAACCAGGA 211

RESULT 6
CA446812/c
LOCUS      673 bp mRNA linear EST 08-NOV-2002
DEFINITION UI-H-EDI-axx-p-03-0-UI.s1 NCI CGAP ED1 Homo sapiens cDNA clone
            UI-H-EDI-axx-p-03-0-UI 3', mRNA sequence.
ACCESSION  CA446812
VERSION     1 GI:24811232
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.

REFERENCE
1 (bases 1 to 673)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1. .673
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="UI-H-EDI-axx-p-03-0-UI"
   /tissue_type="Chondrosarcoma"
   /dev_stage="Adult"
   /lab_host="DH10B (Life Technologies)"
   /clone_lib="NCI CGAP ED1"
   /note="Organ: Left Pubic Bone; Vector: pT7T3-Pac
   (Pharmacia) with a modified polylinker; Site 1: EcoR I;
   Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library
   containing the following tissue(s): Chondrosarcoma cell
   line C55. The library was constructed according to
   Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
   1996. First strand cDNA synthesis was primed with an
   oligo-dT primer containing a Not I site. Double stranded
   cDNA was ligated to an EcoR I adaptor, digested with Not
   I, and cloned directionally into pT7T3-Pac vector. The
   oligonucleotide used to prime the synthesis of
   first-strand cDNA contains a library tag sequence that is

```

located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_TISSUE=chondrosarcoma
TAG_LIB=UI-H-ED1
TAG_SEQ=CGTCAAGGCT

ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 673;
Best Local Similarity 100.0%; Pred. No. 3.7e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
|||||
670 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 620

Db

RESULT 7

BM041386/c
LOCUS 773 bp mRNA linear EST 07-NOV-2001
DEFINITION 603614446T1 NIH_MGC_108 Homo sapiens cDNA clone IMAGE:5420314 3',
mRNA sequence.

ACCESSION BM041386

VERSION BM041386.1 GI:16770653

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 773)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM1874 row: 1 column: 11

High quality sequence start: 26

High quality sequence stop: 718.

FEATURES

source

1..773
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5420314"
/tissue_type="Wilms' tumor, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 108"
/notes="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 51; DB 3; Length 773;
Best Local Similarity 100.0%; Pred. No. 3.8e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
|||||
611 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 561

Db

RESULT 8

AL576621/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 864)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 16, 2001 this sequence version replaced gi:31314903.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Crenieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

8342.f

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CSODI078AG06NP1&c=8342.f>.

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI078YM11"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 51; DB 1; Length 864;
Best Local Similarity 100.0%; Pred. No. 3.8e-07; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 1 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 51
|||||

Db 616 ACAGAACTGGCAAGGCAAGAGGTCACCTAGGGGCTCTGTCAACCCAGGA 566
|||||

RESULT 9

B50582/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 307)

Kim, U.-J., Adams, M.D. and Simon, M.I.

Determination of clone end sequences of human Bacterial Artificial

Chromosomes

Unpublished (1997)

Contact: Ung-Jin Kim

CalTech Genome Research Lab

AL576621 864 bp mRNA linear EST 07-APR-2004
AL576621 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI078YM11 3-PRIME, mRNA sequence.

ACCESSION AL576621

VERSION AL576621

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 864)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 16, 2001 this sequence version replaced gi:31314903.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Crenieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

8342.f

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CSODI078AG06NP1&c=8342.f>.

Location/Qualifiers

1..864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI078YM11"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

B50582 307 bp DNA linear GSS 20-JUN-1998
CIT-HSP-494012.TV CIT-HSP Homo sapiens genomic clone 494012,
genomic survey sequence.

ACCESSION B50582

VERSION B50582.1 GI:2602819

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 307)

Kim, U.-J., Adams, M.D. and Simon, M.I.

Determination of clone end sequences of human Bacterial Artificial

Chromosomes

Unpublished (1997)

Contact: Ung-Jin Kim

CalTech Genome Research Lab

ORIGIN

Query Match 73.3%; Score 37.4; DB 2; Length 855;
 Best Local Similarity 96.1%; Pred. No. 0.024; Indels 1; Gaps 1;
 Matches 49; Conservative 0; Mismatches 1;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGGA 51
 |||||
 Db 572 ACAGAACTGGCAAGAGGCAAGAGGTCACTGA-GGGCTCTGTCAACCAGGA 621
 |||||

RESULT 12

AA324242
 LOCUS
 DEFINITION
 EST27116 Cerebellum II Homo sapiens cDNA 5' end similar to EST
 containing Alu repeat, mRNA sequence.

ACCESSION
 AA324242

VERSION
 AA324242.1 GI:1976559

KEYWORDS
 EST.

SOURCE
 Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.

REFERENCE
 1 (bases 1 to 500)

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
 Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Falancques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M. and Venter,J.C.

Initial assesment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

7566098

JOURNAL

PUBMED

COMMENT

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/cdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1. 500

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):124792"

/db_xref="taxon:9606"

/tissue_type="cerebellum"

/dev_stage="adult"

/clone_lib="Cerebellum II"

/notes="Organ: brain; Vector: pBluescript SK-; Site_1:

EcORI; Site_2: XhoI"

ORIGIN

Query Match 55.7%; Score 28.4; DB 1; Length 500;
 Best Local Similarity 80.0%; Pred. No. 34;
 Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY

8 TGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACC 47
 |||||

Db

366 TGGCAAGAGGCAAGAGGTCACTGAGGGCTTTTNNACC 405
 |||||

RESULT 13

LOCUS

CV358408
 DEFINITION
 PMO-AN0087-270101-008-g09 AN0087 Homo sapiens cDNA, mRNA sequence.

ACCESSION
 CV358408

VERSION
 CV358408.1 GI:52708463

KEYWORDS
 EST.

SOURCE
 Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.

REFERENCE
 1 (bases 1 to 412)

AUTHORS

Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brumstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

JOURNAL

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. <http://www.ludwig.org.br>.

FEATURES

source

1. 412
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="AN0087"
 /note="Organ: amnion normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 54.5%; Score 27.8; DB 7; Length 412;

Best Local Similarity 74.5%; Pred. No. 53;

Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY

4 GAACGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGG 50
 |||||

Db

46 GAACGGCAAGAGGCTAAAGGTCACCGTGGTCTCTGGCCCTCCAGG 92
 |||||

RESULT 14

LOCUS

BE385462
 DEFINITION
 601276049F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3616813 5',

mRNA sequence.

BE385462

BE385462.1 GI:9330827

KEYWORDS
 EST.

SOURCE
 Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 680)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM283 row: b column: 14
High quality sequence stop: 617.
Location/Qualifiers
1..680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3616813"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 20"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

FEATURES
source
Query Match 52.9%; Score 27; DB 2; Length 680;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 1 ACAGACTGGCAAGAGGCAAGAGGTCTACTGAGGGCCTCTGTCTCACCAGGA 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 ACAGAGGTCTCAAGACCCAGAGGACTCTGGGAGGCGCCTGCCACCCAGGA 317

RESULT 15
BI869166/c
LOCUS
DEFINITION BI869166 698 bp mRNA linear EST 11-OCT-2001
603395496F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405446 5',
mRNA sequence.
BI869166
ACCESSION BI869166.1 GI:16042826
VERSION BI869166.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 698)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12033 row: p column: 23

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 698)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12033 row: p column: 23

FEATURES
source
High quality sequence stop: 539.
Location/Qualifiers
1..698
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405446"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sall; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 52.2%; Score 26.6; DB 3; Length 698;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 3 AGAACTGGCAAGAGGCAAGAGGTCTACTGAGGGCCTCTGTCTCACCAGGA 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 650 AGAAGGGTAAATAGTGAAGGGGAGCTGAGGCTTGAGTCAACCAGGA 602

Search completed: February 7, 2006, 14:53:04
Job time : 204.677 secs

This Page Blank (uspio)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:22:44 ; Search time 31.0376 Seconds
(without alignments)
10951.209 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236
Perfect score: 51
Sequence: 1 acagaactggcaaggca.....agggctctgtcaccaggga 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	365	3 AAC98658	AAC98658 Human col
2	51	100.0	1210	13 ACN41203	ACN41203 Human dia
3	51	100.0	1290	13 ACN41200	ACN41200 Human dia
4	51	100.0	1305	13 ACN41199	ACN41199 Human dia
5	51	100.0	1777	14 ADW12906	ADW12906 Human SMA
6	51	100.0	1837	14 ADW12904	ADW12904 Human SMA
7	50	98.0	13808	8 AAL36986	AAL36986 Human mus
8	50	98.0	13808	8 ABX59974	ABX59974 cDNA enco
9	50	98.0	13808	12 ADJ30724	ADJ30724 Human mus
10	25.6	50.2	99916	6 ADI03931	ADI03931 Human enz
11	25.6	50.2	116297	12 ADQ97587	ADQ97587 Human can
12	24.8	48.6	992	5 ABV21964	ABV21964 Human pro
13	24.8	48.6	992	5 ABV27796	ABV27796 Human pro
14	24.4	47.8	331	4 AAL36705	AAL36705 Human mus
15	24.4	47.8	331	4 AAL36704	AAL36704 Human mus
16	24.4	47.8	331	8 ABX59692	ABX59692 cDNA enco
17	24.4	47.8	331	8 ABX59693	ABX59693 cDNA enco
18	24.4	47.8	331	12 ADJ30443	ADJ30443 Human mus
19	24.4	47.8	331	12 ADJ30442	ADJ30442 Human mus

C	20	24.4	47.8	1023	14	ACL73128	ACL73128 M. xanthu
C	21	24.4	47.8	13346	14	ACL64626	ACL64626 M. xanthu
C	22	23.8	46.7	1008	3	AAC74422	AAC74422 Human sec
C	23	23.8	46.7	1008	8	ABZ73645	ABZ73645 Secreted
C	24	23.8	46.7	1008	8	ADA98137	ADA98137 Human sec
C	25	23.8	46.7	1008	10	ABZ67239	ABZ67239 Human sec
C	26	23.8	46.7	1043	3	AAC93331	AAC93331 Human sec
C	27	23.8	46.7	1312	4	AAS35023	AAS35023 Human col
C	28	23.8	46.7	2140	13	ADU82622	ADU82622 Human MDD
C	29	23.8	46.7	3863	4	AAS34888	AAS34888 cDNA enco
C	30	23.8	46.7	3863	10	ADC46046	ADC46046 Human neo
C	31	23.8	46.7	247682	12	ADL08109	ADL08109 Human gen
C	32	23.6	46.3	303	10	ADI22195	ADI22195 Rat liver
C	33	23.6	46.3	506	12	ACH73473	ACH73473 Human gen
C	34	23.6	46.3	7862	12	ADQ22161	ADQ22161 Human sof
C	35	23.4	45.9	2932	8	ADA41546	ADA41546 Human sec
C	36	23.4	45.9	2932	10	ADA57677	ADA57677 BAC fragm
C	37	23.4	45.9	17173	8	ABZ80234	ABZ80234 Mouse tra
C	38	23.2	45.5	1756	11	ADM01710	ADM01710 Human CDN
C	39	23.2	45.5	3089	10	ADB62760	ADB62760 Human CDN
C	40	23.2	45.5	3356	13	ADR08430	ADR08430 Full leng
C	41	23.2	45.5	11706	5	ABA20479	ABA20479 Human ner
C	42	23.2	45.5	28564	10	ADD47140	ADD47140 Human gen
C	43	23.2	45.5	34875	10	ADD14677	ADD14677 Human src
C	44	23.2	45.5	34875	13	ADR52965	ADR52965 Drug ther
C	45	23	45.1	556	4	AAL18474	AAL18474 Probe #84

ALIGNMENTS

RESULT 1
AAC98658
ID AAC98658 standard; cDNA; 365 BP.

AC AAC98658;
XX
XX 09-MAR-2001 (first entry)
XX Human colon cancer antigen nucleotide sequence SEQ ID NO:668.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX identification; cytostatic; cardioactive; neuroprotective; vulnerary;
XX immunomodulatory; muscular; gynaecological; gastrointestinal;
XX nephrotropic; anti-infective; antibacterial; gene therapy; wound;
XX neural disorder; immune system disorder; muscular disorder;
XX reproductive disorder; gastrointestinal disorder; renal disorder;
XX infectious disease; cardiovascular disorder; ss.

OS Homo sapiens.
XX WO200055351-A1.
XX
XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005883.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587534/55.

XX P-PSDB; AAB53901.

XX Colon cancer associated gene sequences, referred to as colon cancer
XX antigens, useful for the treatment, prevention, and diagnosis of colon
XX disorders such as colon cancer.

XX Claim 1; Page 1200; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,

CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardiactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and
CC vulnary, nephrotropic, antineoplastic and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 365 BP; 54 A; 121 C; 101 G; 79 T; 0 U; 10 Other;

Query Match 100.0%; Score 51; DB 3; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGGA 51
|||||
Db 32 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGGA 82
|||||

RESULT 2
ACN41203
ID ACN41203 standard; cDNA; 1210 BP.
XX
AC ACN41203;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:78.
XX
KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR WPI; 2004-329368/30.
DR P-PSDB; ABM82551.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 1; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC

CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorders, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1210 BP; 248 A; 375 C; 376 G; 211 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 13; Length 1210;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGGA 51
|||||
Db 1024 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGGA 1074
|||||

RESULT 3
ACN41200
ID ACN41200 standard; cDNA; 1290 BP.
XX
AC ACN41200;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:75.
XX
KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR WPI; 2004-329368/30.
DR P-PSDB; ABM82548.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 1; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC

CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1290 BP; 261 A; 402 C; 398 G; 229 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 13; Length 1290;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 51
|||||
DB 1104 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 1154

RESULT 4
ACN41199
ID ACN41199 standard; cDNA; 1305 BP.
XX
AC ACN41199;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:74.
XX
KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX Harthorne TA, Suchorolski MT, Altus CM, Pette SU, Elder LV;
XX Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LH;
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
XX Patury S, Shi X, Suarez CJ;
XX
XX WPI; 2004-329368/30.
XX P-PSDB; ABM82547.

XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX
XX Claim 1; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 2722 sequences defined in the specification. A

CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1305 BP; 273 A; 411 C; 392 G; 229 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 13; Length 1305;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 51
|||||
DB 1119 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCAGGA 1169

RESULT 5
ADM12906
ID ADM12906 standard; cDNA; 1777 BP.
XX
AC ADM12906;
XX
DT 07-APR-2005 (first entry)
XX
DE Human SMAPK3V4 variant gene.
XX
KW DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW gene; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 12..1091
FT /tag= a
FT /product= "SMAPK3V4 protein"
XX
XX US2005013817-A1.
XX
XX 20-JAN-2005.
XX
XX 18-JUL-2003; 2003US-00623108.
XX
XX 18-JUL-2003; 2003US-00623108.
XX
XX (DAIK/) DAI K.
XX
XX Dai K;
XX
XX WPI; 2005-080923/09.
XX P-PSDB; ADM12907.

XX New isolated SMAPK3 polypeptides and nucleic acids, useful for diagnosing
XX diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene
XX in a mammal.
XX
XX Claim 4; SEQ ID NO 7; 55pp; English.

XX The invention relates to novel isolated polypeptide (I) comprising an
XX amino acid sequence selected from sequences comprising 335 or 359 amino
XX acids (ADM12901 or ADM12903) or its fragments. The polypeptides and
XX polynucleotides are useful for diagnosing diseases, e.g. cancers,
XX associated with the deficiency of the SMAPK3 gene in a mammal. The
XX fragments of the polypeptides and polynucleotides can also be used as

CC primers or probes. This sequence corresponds to the cDNA encoding the
CC SNAPK3V1 variant protein.

SQ Sequence 1777 BP; 355 A; 575 C; 507 G; 340 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 14; Length 1777;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCTCTGTCAACCCAGGA 51
Db 1126 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCTCTGTCAACCCAGGA 1176

RESULT 6
ADW12904
ID ADW12904 standard; cDNA; 1837 BP.

XX AC ADW12904;
XX 07-APR-2005 (first entry)
XX Human SNAPK3V3 variant gene.
XX DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
XX gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 12..1151
XX /*tag= a
XX /product= "SNAPK3V3 protein"

XX US2005013817-A1.
XX 20-JAN-2005.
XX 18-JUL-2003; 2003US-00623108.
XX 18-JUL-2003; 2003US-00623108.
XX (DAIK/) DAI K.
XX Dai K;
XX WPI; 2005-080923/09.
XX P-PSDB; ADW12905.
XX New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing
XX diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene
XX in a mammal.

XX Claim 4; SEQ ID NO 5; 55pp; English.
XX The invention relates to novel isolated polypeptide (I) comprising an
XX amino acid sequence selected from sequences comprising 335 or 359 amino
XX acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
XX polynucleotides are useful for diagnosing diseases, e.g. cancers,
XX associated with the deficiency of the SNAPK3 gene in a mammal. The
XX fragments of the polypeptides and polynucleotides can also be used as
XX primers or probes. This sequence corresponds to the cDNA encoding the
XX SNAPK3V1 variant protein.

SQ Sequence 1837 BP; 369 A; 593 C; 526 G; 349 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 14; Length 1837;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCTCTGTCAACCCAGGA 51
Db 1186 ACAGAACTGGCAAGAGGCAAGAGGTCACCTAGGGCCTCTGTCAACCCAGGA 1236

RESULT 7

AAL36986
ID AAL36986 standard; DNA; 13808 BP.

XX AAL36986;
XX 08-JAN-2002 (first entry)
XX Human musculoskeletal system related polynucleotide SEQ ID NO 3351.
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
XX vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ds.
XX Homo sapiens.
XX WO200155367-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001338.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214866P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216800P.
XX 11-JUL-2000; 2000US-0217487P.
XX 14-JUL-2000; 2000US-0217496P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 18-AUG-2000; 2000US-0225759P.
XX 22-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226688P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 05-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 06-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232337P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 12-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
XX cancers and also for testing and detection e.g. diagnosis.
XX
XX Example 2; SEQ ID NO 3351; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABB03087-ABB04109) associated with the musculoskeletal system useful for
XX preventing, treating or ameliorating medical conditions e.g. By protein
XX or gene therapy. The genes are isolated from a range of human tissues
XX disclosed in the specification. The nucleic acids, proteins, antibodies
XX and (ant)agonists are useful in the diagnosis, treatment and prevention
XX of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
XX lung, or urogenital; (b) immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such as
XX myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
XX cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
XX bacterial, fungal and parasitic infections. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 13808 BP; 2990 A; 3953 C; 3656 G; 3209 T; 0 U; 0 Other;
Query Match 98.0%; Score 50; DB 4; Length 13808;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACAGAACTGGCAAGAGGCAAGAGGTCAGTACAGGGCCCTCTGTACCCAGG 50
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCAGTACAGGGCCCTCTGTACCCAGG 6424
RESULT 8
ABX59974
ID ABX59974 standard; cDNA; 13808 BP.
XX
XX AC ABX59974;
XX
XX 26-FEB-2003 (first entry)
XX cDNA encoding novel human musculoskeletal system antigen #2318.
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX

XX US2004009488-A1.
XX 15-JAN-2004.
XX 13-SEP-2002; 2002US-00242515.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218230P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
XX 26-SEP-2000; 2000US-0235484P.
XX 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764877.
XX (HUMA-) HUMAN GENOME SCI INC.
XX FA

XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2004-090458/09.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g., cancer of
PT musculoskeletal tissues or osteoporosis.
XX Disclosure; SEQ ID NO 3351; 289pp; English.
XX The invention relates to a novel isolated musculoskeletal system-
CC associated nucleic acid molecule. The nucleic acid of the invention
CC demonstrates cytostatic and osteopathic activities and may be useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition such as cancer of the musculoskeletal tissues or osteoporosis,
CC possibly via gene therapy or vaccine production. The current sequence is
CC that of the human musculoskeletal system-associated genomic DNA of the
CC invention. The current sequence is not shown within the specification per
CC se but is available on the USPTO web-site
CC http://seqdata.uspto.gov/sequence.html?DocID=20040009488.
XX
SQ Sequence 13808 BP; 2990 A; 3953 C; 3656 G; 3209 T; 0 U; 0 Other;
Query Match 98.0%; Score 50; DB 12; Length 13808;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 50
Db 6375 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAGG 6424
RESULT 10
AD103931
ID AD103931 standard; DNA; 99916 BP.
XX
AC AD103931;
XX
DT 22-APR-2004 (first entry)
XX
DE Human enzyme protein encoding genomic DNA.
XX
KW Enzyme protein; drug screening; disease diagnosis; human; gene therapy;
KW chromosome 6; gene; db; glucuronyltransferase.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FT 2614..96915
CDS /*tag= b
FT /*product= "enzyme protein"
FT /*note= "contains introns"
FT 2614..3204
FT /*tag= a
FT /*tag= 1
FT /*number= 1
FT 3205..64312
FT /*tag= c
FT /*number= 1
FT 64313..64457
FT /*tag= d
FT /*number= 2
FT 64458..96602
FT /*tag= e
FT /*number= 2
FT 96603..96751
FT /*tag= f
FT /*number= 3
FT 96752..96831
FT /*tag= g
FT /*number= 3
FT 96832..96915
FT /*tag= h
FT

FT /number= 4
XX WO200268657-A2.
PN 06-SEP-2002.
XX 08-FEB-2002; 2002WO-US003623.
PF 26-FEB-2001; 2001US-0270871P.
PR 26-MAR-2001; 2001US-00816095.
XX (PEXE) PE CORP NY.
PA Gan W, Yan C, Merkulov GV, Di Francesco V, Beasley EM;
XX WPI; 2002-713380/77.
DR P-PSDB; ADI03930.
XX New human enzyme proteins, useful for treating or diagnosing disorders
PT associated with abnormal expression of the protein, in drug screening
PT assays and pharmacogenomic analysis.
XX Claim 4; SEQ ID NO 3; 127pp; English.
XX The invention relates to a novel isolated enzyme protein and encoding
CC polynucleotides. The protein shows a high degree of similarity to a
CC glucuronyltransferase cloned from a rabbit brain cDNA library. The
CC peptides and nucleic acid molecules are useful as models for the
CC development of human therapeutic targets, aid in the identification of
CC therapeutic proteins, and serve as targets for the development of human
CC therapeutic agents. The peptide may be used in drug screening assays, in
CC assays to determine the biological activity of the protein, to raise
CC antibodies or to elicit another immune response, as a reagent in assays
CC designed to quantitatively determine levels of the protein in biological
CC fluids, or as markers for tissues in which the corresponding protein is
CC preferentially expressed. The human enzyme protein is also useful for
CC diagnosing a disease, predisposition to a disease, or treating a disorder
CC characterized by an absence of, inappropriate or unwanted expression of
CC the protein. The antibodies are useful in pharmacogenomic analysis, for
CC inhibiting protein function, or for tissue typing. The nucleic acid
CC molecules are useful as probes, primers, chemical intermediates, or in
CC encoding genomic DNA.
XX Sequence 99916 BP; 27736 A; 18701 C; 19334 G; 29032 T; 0 U; 5113 Other;
SQ
Query Match 50.2%; Score 25.6; DB 6; Length 99916;
Best Local Similarity 70.8%; Pred. No. 40;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTCAACCCAG 49
Db 1819 CAGACCCGCGGAGAGCAAGGGGTCCTCCGCGCTCTCCTTCGCAG 1866
RESULT 11
ADQ97587
ID ADQ97587 standard; DNA; 116297 BP.
XX
AC ADQ97587;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human cancer associated sequence HD10-009, SEQ ID 564.
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX Homo sapiens.
XX WO2004060304-A2.
PN 22-JUL-2004.
XX


```
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
XX cancers and also for testing and detection e.g. diagnosis.
XX
XX Example 2; SEQ ID NO 3070; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABB03087-ABB04109) associated with the musculoskeletal system useful for
XX preventing, treating or ameliorating medical conditions e.g. by protein
XX or gene therapy. The genes are isolated from a range of human tissues
XX disclosed in the specification. The nucleic acids, proteins, antibodies
XX and (ant)agonists are useful in the diagnosis, treatment and prevention
XX of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
XX lung, or urogenital; (b) immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such as
XX myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
XX cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
XX bacterial, fungal and parasitic infections. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 331 BP; 87 A; 83 C; 78 G; 83 T; 0 U; 0 Other;
XX
XX Query Match 47.8%; Score 24.4; DB 4; Length 331;
XX Best Local Similarity 73.8%; Pred. NO. 32;
XX Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX 2 CAGAACTGGCAAGAGGCAAGGTCACCTAGTGGGCTCTGTC 43
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX DB 244 CAGGACTGACACAGGGGTACAGGCCACTGTGGGAATCTATC 285
```

RESULT 15

AAL36704

ID AAL36704 standard; DNA; 331 BP.

XX

AC AAL36704;

XX

DT 08-JAN-2002 (first entry)

XX

DE Human musculoskeletal system related polynucleotide SEQ ID NO 3069.

XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.

XX

OS Homo sapiens.

XX

PN WO200155367-A1.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001338.

XX

XX 31-JAN-2000; 2000US-0179065P.

XX

XX 04-FEB-2000; 2000US-0180628P.

XX

XX 24-FEB-2000; 2000US-0184664P.

XX

XX 02-MAR-2000; 2000US-0186350P.

XX

XX 16-MAR-2000; 2000US-0189874P.

XX

XX 17-MAR-2000; 2000US-0190076P.

XX

XX 18-APR-2000; 2000US-0198123P.

XX

XX 19-MAY-2000; 2000US-0205515P.

XX

XX 07-JUN-2000; 2000US-0209467P.

XX

XX 28-JUN-2000; 2000US-0214886P.

XX

XX 30-JUN-2000; 2000US-0215135P.

XX

XX 07-JUL-2000; 2000US-0216647P.

XX

XX 07-JUL-2000; 2000US-0216880P.

XX

XX 11-JUL-2000; 2000US-0217487P.

XX

XX 14-JUL-2000; 2000US-0217496P.

XX

XX 28-JUL-2000; 2000US-0218290P.

XX

XX 26-JUL-2000; 2000US-0220963P.

XX

XX 14-AUG-2000; 2000US-0220964P.

XX

XX 14-AUG-2000; 2000US-0224518P.

XX

XX 14-AUG-2000; 2000US-0225213P.

XX

XX 14-AUG-2000; 2000US-0225214P.

XX

XX 14-AUG-2000; 2000US-0225466P.

XX

XX 14-AUG-2000; 2000US-0225467P.

XX

XX 14-AUG-2000; 2000US-0225757P.

XX

XX 14-AUG-2000; 2000US-0225758P.

XX

XX 18-AUG-2000; 2000US-0226279P.

XX

XX 22-AUG-2000; 2000US-0226681P.

XX

XX 22-AUG-2000; 2000US-0226868P.

XX

XX 22-AUG-2000; 2000US-0227182P.

XX

XX 23-AUG-2000; 2000US-0227009P.

XX

XX 30-AUG-2000; 2000US-0228924P.

XX

XX 01-SEP-2000; 2000US-0229287P.

XX

XX 01-SEP-2000; 2000US-0229343P.

XX

XX 01-SEP-2000; 2000US-0229344P.

XX

XX 01-SEP-2000; 2000US-0229345P.

XX

XX 05-SEP-2000; 2000US-0229509P.

XX

XX 05-SEP-2000; 2000US-0229513P.

XX

XX 06-SEP-2000; 2000US-0230437P.

XX

XX 06-SEP-2000; 2000US-0230438P.

XX

XX 08-SEP-2000; 2000US-0231242P.

XX

XX 08-SEP-2000; 2000US-0231243P.

XX

XX 08-SEP-2000; 2000US-0231244P.

XX

XX 08-SEP-2000; 2000US-0231245P.

XX

PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-451937/48.
XX Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including musculoskeletal
PT cancers and also for testing and detection e.g. diagnosis.
XX Example 2; SEQ ID NO 3069; 781pp + Sequence Listing; English.
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABR03087-ABR04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 331 BP; 87 A; 83 C; 78 G; 83 T; 0 U; 0 Other;

Query Match 47.8%; Score 24.4; DB 4; Length 331;

Best Local Similarity 73.8%; Pred. No. 32;

Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTAGGGGCTCTGTC 43

DB 244 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 285

Search completed: February 7, 2006, 12:49:05
Job time : 34.0376 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:27:59 ; Search time 275.097 Seconds
(without alignments)
10538.142 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236
Perfect score: 51
Sequence: 1 acsgaactgcgaaggca.....agggcctctgtcaccaggga 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	50	98.0	187383	14 AC138954	AC138954 Homo sapi
C 2	50	98.0	192943	8 AC012645	AC012645 Homo sapi
C 3	29.6	58.0	249737	14 AC094359	AC094359 Rattus no
C 4	29.6	58.0	251990	14 AC128457	AC128457 Rattus no
5	27.6	54.1	189212	14 AC109240	AC109240 Mus muscu
6	27.6	54.1	230186	9 AL731678	AL731678 Mouse DNA
7	27.6	54.1	239692	14 AC158791	AC158791 Mus muscu
C 8	27.6	54.1	242194	14 AC091314	AC091314 Mus muscu
9	27.6	54.1	329223	14 AC109504	AC109504 Mus muscu
10	27.4	53.7	173029	9 AC132271	AC132271 Mus muscu
11	27.4	53.7	198890	9 AC131679	AC131679 Mus muscu
C 12	26.6	52.2	231980	14 AC094305	AC094305 Rattus no
C 13	26.6	52.2	261198	14 AC106528	AC106528 Rattus no
14	26.6	52.2	282320	14 AC149647	AC149647 Bos tauru
15	26.2	51.4	136581	8 AC008459	AC008459 Homo sapi
C 16	26.2	51.4	137324	14 AC010459	AC010459 Homo sapi
17	26.2	51.4	143976	14 AC141543	AC141543 Rattus no
C 18	26.2	51.4	184015	9 AC125458	AC125458 Mus muscu

19	26.2	51.4	252996	14 AC130153	AC130153 Rattus no
20	26.2	51.4	345997	1 BX569692	BX569692 Synechoco
21	26	51.0	87746	8 AC109826	AC109826 Homo sapi
22	25.8	50.6	49004	8 AL160288	AL160288 Human DNA
C 23	25.8	50.6	56702	8 AY923052	AY923052 Homo sapi
24	25.8	50.6	150236	14 AC023292	AC023292 Homo sapi
25	25.6	50.2	689	8 HSA339409	HSA339409 Homo sapi
26	25.6	50.2	703	8 HSA339162	HSA339162 Homo sapi
27	25.6	50.2	720	8 HSA324773	HSA324773 Homo sapi
C 28	25.6	50.2	92019	8 AL450320	AL450320 Human DNA
29	25.6	50.2	99916	6 AR438683	AR438683 Sequence
30	25.6	50.2	99916	6 AX925386	AX925386 Sequence
C 31	25.6	50.2	179691	4 BX296515	BX296515 Pig DNA s
32	25.6	50.2	184512	8 AC079804	AC079804 Homo sapi
C 33	25.6	50.2	191318	14 AC020670	AC020670 Homo sapi
C 34	25.6	50.2	191937	8 AC100823	AC100823 Homo sapi
C 35	25.4	49.8	153028	9 AL928739	AL928739 Mouse DNA
C 36	25.4	49.8	186647	9 AC129302	AC129302 Mus muscu
C 37	25.4	49.8	187065	9 AC113014	AC113014 Mus muscu
C 38	25.4	49.8	190780	9 AC153508	AC153508 Mus muscu
C 39	25	49.0	142943	8 AC099405	AC099405 Papio anu
C 40	25	49.0	166513	14 AC073787	AC073787 Mus muscu
41	25	49.0	180223	9 AL589692	AL589692 Mouse DNA
C 42	25	49.0	201676	14 AC105367	AC105367 Sus scrofa
43	25	49.0	208651	9 AC090493	AC090493 Genomic s
C 44	25	49.0	209182	9 AC149052	AC149052 Mus muscu
45	25	49.0	216022	14 AC156179	AC156179 Bos tauru

ALIGNMENTS

RESULT 1
AC138954
LOCUS
DEFINITION
AC138954
Homo sapiens chromosome 16 clone RP11-828J20, WORKING DRAFT
SEQUENCE, 5 unordered pieces.
187383 bp DNA linear HTG 21-JAN-2003

ACCESSION
AC138954.1 GI:27805366
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 187383)
DOE Joint Genome Institute.
AUTHORS
Sequencing of Human Chromosome 16
TITLE
Unpublished
JOURNAL
2 (bases 1 to 187383)
DOE Joint Genome Institute.
REFERENCE
Direct Submission
AUTHORS
Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint
TITLE
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
JOURNAL
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1612692
Center clone name: RPCI-11_828J20

Summary Statistics
Consensus quality: 181858 bases at least Q40
Consensus quality: 184513 bases at least Q30
Consensus quality: 185456 bases at least Q20
Estimated insert size: 180000; agarose-fp estimation
Estimated insert size: 186983; sum-of-contigs estimation
Quality coverage: 8.77 in Q20 bases; agarose-fp estimation
Quality coverage: 8.44 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1225: contig of 1225 bp in length
 * 1226 1325: gap of unknown length
 * 1326 34247: contig of 32922 bp in length
 * 34248 71258: gap of unknown length
 * 71259 71358: contig of 36911 bp in length
 * 71359 117554: gap of unknown length
 * 117555 117654: contig of 46196 bp in length
 * 117655 187383: gap of unknown length
 * 187384 69729: contig of 69729 bp in length.

FEATURES

source

1. .187383
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-828J20"
 /clone_lib="RP11-828J20" human BAC library 11"
 1226. .1325
 /estimated_length=unknown
 34248. .34347
 /estimated_length=unknown
 71259. .71358
 /estimated_length=unknown
 117555. .117654
 /estimated_length=unknown

ORIGIN

Query Match 98.0%; Score 50; DB 14; Length 187383;
 Best Local Similarity 100.0%; Pred. No. 5.2e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 50
 |||||
 Db 41734 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 41783

RESULT 2

AC012645/c 192943 bp DNA linear PRI 07-JAN-2004
 LOCUS Homo sapiens chromosome 16 clone RP11-455F5, complete sequence.
 AC012645
 AC012645.7 GI:40714622

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE

1 (bases 1 to 192943)
 DOE Joint Genome Institute, Stanford Human Genome Center and Los
 Alamos National Laboratory.

AUTHORS

Direct Submission

TITLE

Unpublished

JOURNAL

DOE Joint Genome Institute.

REFERENCE

2 (bases 1 to 192943)

AUTHORS

DOE Joint Genome Institute.

TITLE

Direct Submission

JOURNAL

Submitted (31-OCT-1998) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE

3 (bases 1 to 192943)

AUTHORS

DOE Joint Genome Institute.

TITLE

Direct Submission

JOURNAL

Submitted (07-JAN-2004) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA

COMMENT

On Jan 7, 2004 this sequence version replaced gi:19172834.

Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center and Los Alamos
 National Laboratory

www.shgc.stanford.edu

Quality: Phrap Quality >=40 100% of Sequence;

Estimated Total Number of Errors is 0.

Location/Qualifiers

1. .192943

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="16"

/clone="RP11-455F5"

ORIGIN

Query Match 98.0%; Score 50; DB 8; Length 192943;

Best Local Similarity 100.0%; Pred. No. 5.2e-09;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 50
 |||||
 Db 53010 ACAGAACTGGCAAGAGGCAAGAGGTCACCTGAGGGCTCTGTGTCACCCAGG 52961

RESULT 3

AC094359/c 249737 bp DNA linear HTG 09-MAY-2003

LOCUS Rattus norvegicus clone CH230-3C19, WORKING DRAFT SEQUENCE, 3
 unordered pieces.

ACCESSION

AC094359.6 GI:30467240

VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 249737)

AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Caesar, H., Center, A.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Y., Chen, Z., Chu, J.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Y., Chen, Y.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoggess, J.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshew, L., Louise, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 249737)
Worley, K.C.

Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 249737)
Baylor Genome Sequencing Consortium.

Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24942531.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GALLI
Center clone name: CH230-3C19
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 236591 bases at least Q40
Consensus quality: 239010 bases at least Q30
Consensus quality: 241077 bases at least Q20
Estimated insert size: 247522; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 89596: contig of 89596 bp in length
* 89597 89696: gap of unknown length

* 89697 248439: contig of 158743 bp in length
* 248440 248539: gap of unknown length
* 248540 249737: contig of 1198 bp in length.

FEATURES
source
1..249737
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3C19"
54592..54643
/note="clone boundary
clone end:SP6
site:ECORI
end_sequence: BH307413"
89597..89696
/estimated_length=unknown
89697..90719
/note="wgs contig"
248440..248539
/estimated_length=unknown

misc_feature
gap
misc_feature
gap

ORIGIN
Query Match 58.0%; Score 29.6; DB 14; Length 249737;
Best Local Similarity 79.5%; Pred. No. 0.98;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCATCTAGGCGCTCTGTCA 44
|||||
Db 222041 ACAGATCTGGCAAGAGGCGAGGTCAGTCTGCTGCTCTTA 221998
|||||

RESULT 4
AC128457/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-85E19, *** SEQUENCING IN PROGRESS
*** 9 unordered pieces.
AC128457
AC128457.2 GI:232467595
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
VERSION
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 251990)
Muzny, D., Marle, Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, Y., Chen, Y., Chen, Y., Chen, Y., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falla, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Harwayne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hollins, B., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollings, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Napua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,

Miloslavjevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Pank, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE JOURNAL REFERENCE AUTHORS

Direct Submission
Unpublished
2 (bases 1 to 251990)
Worley, K.C.

TITLE JOURNAL

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 251990)
Rat Genome Sequencing Consortium.

REFERENCE AUTHORS TITLE JOURNAL

Direct Submission
Submitted (10-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYVR
Center clone name: CH230-85E19

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 190117 bases at least Q40
Consensus quality: 195766 bases at least Q30
Consensus quality: 199840 bases at least Q20
Estimated insert size: 189582; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

* be preserved.
* 1 11689: contig of 11689 bp in length
* 11690 gap of unknown length
* 11789: contig of 224856 bp in length
* 236645: contig of unknown length
* 236745: gap of unknown length
* 241468: contig of 4723 bp in length
* 241568: gap of unknown length
* 241569: contig of 1428 bp in length
* 242997: gap of unknown length
* 243097: contig of 1185 bp in length
* 244281: contig of unknown length
* 244382: gap of unknown length
* 244382: contig of 1133 bp in length
* 245515: gap of unknown length
* 245615: contig of 1118 bp in length
* 246733: gap of unknown length
* 246832: contig of unknown length
* 248947: contig of 2115 bp in length
* 248948 gap of unknown length
* 249048: contig of 2943 bp in length.
FEATURES
Location/Qualifiers
1..251990
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-85E19"
1..1408
/note="wgs end extension
clone end:Sp6"
complement(6981..7869)
/note="clone boundary
clone end:Sp6
site:EcoRI
end sequence:BH304395"
11690..11789
/estimated_length=unknown
11790..12959
/note="wgs_contig"
13461..14737
/note="wgs_contig"
33176..34795
/note="wgs_contig"
37663..39153
/note="wgs_contig"
67847..69250
/note="wgs_contig"
72977..74360
/note="wgs_contig"
131426..132574
/note="wgs_contig"
236646..236745
/estimated_length=unknown
241469..241568
/estimated_length=unknown
242997..243096
/estimated_length=unknown
244282..244381
/estimated_length=unknown
245515..245614
/estimated_length=unknown
246733..246832
/estimated_length=unknown
248948..249047
/estimated_length=unknown
ORIGIN
Query Match 58.0%; Score 29.6; DB 14; Length 251990;
Best Local Similarity 79.5%; Pred.No. 0.98;
Matches 35; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ACAGACTGGCAAGAGGCAAGAGGTCACTGAGGCGCTCTGTCA 44
DB 207842 ACAGATCTGGCAAGAGGCAAGAGGTGACTGCTGCTCTTA 207799

```

RESULT 5
AC109240
LOCUS
DEFINITION
    Mus musculus chromosome 1 clone RP23-445123 map 1, linear HTG 28-JUN-2005
IN PROGRESS ***, 2 ordered pieces.
ACCESSION
    AC109240.29 GI:69268179
VERSION
    HTG; HTGS PHASE2; HTGS FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS
    Mus musculus (house mouse)
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
    Birren,B., Nusbaum,C. and Lander,E.
    1 (bases 1 to 189212)
    Mus musculus chromosome 1, clone RP23-445123
    Unpublished
    2 (bases 1 to 189212)
    Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
    Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
    Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
    Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
    Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
    Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
    Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
    Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
    Kamat,A., Karatas,A., Kellis,C., LaRocque,K., Lamazares,R.,
    Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
    Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
    McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
    Mienga,V., Murphy,T., Naylor,J., Nguyen,J., Nicol,R., Norbu,C.,
    Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
    Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
    Retta,R., Riaback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
    Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
    Sevrans,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
    Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
    Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
    Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
    Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
    Direct Submission
    Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
    Research, 320 Charles Street, Cambridge, MA 02141, USA
    3 (bases 1 to 189212)
    Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
    Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
    Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
    Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
    Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
    Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
    Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
    Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
    Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., Landers,T.,
    Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
    MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
    McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
    Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
    O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
    Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
    Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
    Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
    Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
    Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
    Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
    Zimmer,A. and Zody,M.
    Direct Submission
    Submitted (28-JUN-2005) Broad Institute of MIT and Harvard, 320
    Charles Street, Cambridge, MA 02141, USA
    On Jun 28, 2005 this sequence version replaced gi:66841635.
    All repeats were identified using RepeatMasker:
    Smit, A.F.A. & Green, P. (1996-1997)
    http://ftp.genome.washington.edu/RM/RepeatMasker.html
    ----- Genome Center
    During sequence assembly data is compared from overlapping clones.

```

```

Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L20866
Center clone name: 445_1_23
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 174728: contig of 174728 bp in length
* 174729 174828: gap of unknown length
* 174829 189212: contig of 14384 bp in length.

```

FEATURES

```

source
    1..189212
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10090"
        /chromosome="1"
        /map="1"
        /clone="RP23-445123"
        /clone_lib="RPCI-23 Female Mouse BAC"
        174729..174828
            /estimated_length=unknown

```

ORIGIN

```

Query Match      54.1%; Score 27.6; DB 14; Length 189212;
Best Local Similarity 72.0%; Pred. No. 6.2;
Matches 36; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ACAGACTGCGAAAGAGGCAAGGTCATGAGGGGCTCTGTGTCACCCAGG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132392 ACAGTCCAGCGCAGAGGCTGGAGGTCCACAGAGTCCCTTTTCACACAGG 132441

```

RESULT 6

```

AL731678
LOCUS
DEFINITION
    Mouse DNA sequence from clone RP23-39H12 on chromosome X, complete
    sequence.
ACCESSION
    AL731678
VERSION
    AL731678.17 GI:27368268
KEYWORDS
    HTG.
SOURCE
    Mus musculus (house mouse)
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
    Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
    1 (bases 1 to 230186)
    Heath,P.
    Direct Submission
    Submitted (21-DEC-2002) Wellcome Trust Sanger Institute, Hinxton,
    Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
    humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
    On Dec 23, 2002 this sequence version replaced gi:26788084.
    Sequence from the Mouse Genome Sequencing Consortium whole genome
    shotgun may have been used to confirm this sequence. Sequence data
    from the whole genome shotgun alone has only been used where it has
    a phred quality of at least 30.

```

COMMENT

```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones.


```
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-247C6"
/clone_lib="RPCI-23 Female Mouse BAC"
88411..88510
/estimated_length=unknown
92995..93094
/estimated_length=unknown
110664..110763
/estimated_length=unknown
137282..137381
/estimated_length=unknown
154682..154781
/estimated_length=unknown
157223..157322
/estimated_length=unknown
167806..167905
/estimated_length=unknown

ORIGIN
Query Match      54.1%; Score 27.6; DB 14; Length 239692;
Best Local Similarity 72.0%; Pred. No. 6.3;
Matches 36; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ACAGAACTGGCAAGAGCGAGAGGTCACCTGCTCTGTCCACCAGG 50
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227642 ACAGTCCAGCGCAGAGGCTGGAGGTCACAGAGTCCCTTCACACAGG 227691

RESULT 8
AC091314/c
LOCUS
DEFINITION Mus musculus clone RP23-57B2, WORKING DRAFT SEQUENCE, 17 unordered
            pieces.
AC091314
VERSION AC091314.3 GI:22325114
KEYWORDS HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Birren,B., Nusbaum,C. and Lander,B.
TITLE Unpublished
JOURNAL
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
            Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
            Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K.,
            Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
            Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRoque,K.,
            Lanazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
            McSwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
            Severy,P., Sougnuez,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (13-APR-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
```

```

* 43026 47226: contig of 4201 bp in length
* 47227 47326: gap of 100 bp
* 47327 57802: contig of 10476 bp in length
* 57803 57902: gap of 100 bp
* 57903 72870: contig of 14968 bp in length
* 72871 72970: gap of 100 bp
* 72971 89540: contig of 16570 bp in length
* 89541 89640: gap of 100 bp
* 89641 162712: contig of 73072 bp in length
* 162713 162812: gap of 100 bp
* 162813 185801: contig of 22989 bp in length
* 185802 185901: gap of 100 bp
* 185902 220592: contig of 34691 bp in length
* 220593 220692: gap of 100 bp
* 220693 242194: contig of 21502 bp in length.
FEATURES
    source
        1..242194
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="RP23-5782"
            /clone_lib="RPC1-23 Female Mouse BAC"
            1..22307
                /note="assembly_fragment"
                clone_end:SP6
                vector_side:left
            22308..22407
                /estimated_length=100
            22408..23603
                /note="assembly_fragment"
            23604..23703
                /estimated_length=100
            23704..26091
                /note="assembly_fragment"
            26092..26191
                /estimated_length=100
            26192..28560
                /note="assembly_fragment"
            28561..28660
                /estimated_length=100
            28661..31163
                /note="assembly_fragment"
            31164..31263
                /estimated_length=100
            31264..32354
                /note="assembly_fragment"
            32355..32454
                /estimated_length=100
            32455..34755
                /note="assembly_fragment"
            34756..34855
                /estimated_length=100
            34856..37785
                /note="assembly_fragment"
            37786..37885
                /estimated_length=100
            37886..42925
                /note="assembly_fragment"
            42926..43025
                /estimated_length=100
            43026..47226
                /note="assembly_fragment"
            47227..47326
                /estimated_length=100
            47327..57802
                /note="assembly_fragment"
            57803..57902
                /estimated_length=100
            57903..72870
                /note="assembly_fragment"
            72871..72970
                /estimated_length=100
            72971..89540

```

```

/note="assembly_fragment"
89541..89640
/estimated_length=100
89641..162712
/note="assembly_fragment"
162713..162812
/estimated_length=100
162813..185801
/note="assembly_fragment"
185802..185901
/estimated_length=100
185902..220592
/note="assembly_fragment"
220593..220692
/estimated_length=100

Query Match      54.1%; Score 27.6; DB 14; Length 242194;
Best Local Similarity 78.6%; Pred. No. 6.3;
Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      8 TGGCAAGAGGCAAGAGTCACTGAGGGCCCTGTGTCCACCAG 49
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      106371 TGGCCACAGGCCAACAGTAAGTACTGAGGCCCTGTGTCCACCAG 106330

RESULT 9
AC109504      329223 bp      DNA      linear      HTG 13-MAR-2004
LOCUS      Mus musculus chromosome 13 clone RP23-184C21 map 13, WORKING DRAFT
DEFINITION      SEQUENCE, 20 unordered pieces.
ACCESSION      AC109504
VERSION      AC109504.5 GI:45430118
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 329223)
Birren,B., Nussbaum,C. and Lander,E.
Mus musculus chromosome 13, clone RP23-184C21
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 329223)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazarot,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (04-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 329223)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavkiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,

```

Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 13, 2004 this sequence version replaced gi:20389429.
All repeats were identified using RepeatMasker:
Snit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL COMMENT

FEATURES
Location/Qualifiers
source
1. .329223 /organism="Mus musculus"
166896: contig of 9572 bp in length /mol_type="genomic DNA"
16697: gap of 100 bp /db_xref="taxon:10090"
222799: contig of 55803 bp in length /chromosome="13"
222800: gap of 100 bp /map="13"
222800: contig of 6670 bp in length /clone="RP23-184C21"
222800: gap of 100 bp /clone_lib="RPC1-23 Female Mouse BAC"
222800: contig of 10630 bp in length 1. .1913
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 47641 bp in length 1914. .2013
222800: gap of 100 bp /estimated_length=100
222800: contig of 4407 bp in length /notes="assembly_fragment"
222800: gap of 100 bp 27037. .27136
222800: contig of 13613 bp in length 27137. .31366
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 3790 bp in length 31367. .31466
222800: gap of 100 bp /estimated_length=100
222800: contig of 166896 bp in length 31467. .85408
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 18973 bp in length 85409. .85508
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 85509. .89621
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 89622. .89721
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 89722. .104564
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 104565. .104664
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 104665. .112440
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 112441. .112540
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 112541. .131698
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 131699. .131798
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 131799. .142284
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 142285. .142384
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 142385. .151005
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 151006. .151105
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 151106. .157224
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 157225. .157324
222800: gap of 100 bp /estimated_length=100
222800: contig of 166897 bp in length 157325. .166896
222800: gap of 100 bp /notes="assembly_fragment"
222800: contig of 166897 bp in length 166897. .166996
222800: gap of 100 bp

NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1913: contig of 1913 bp in length
1914 2013: gap of 100 bp
2014 27036: contig of 25023 bp in length
27037 27136: gap of 100 bp
27137 31366: contig of 4230 bp in length
31367 31466: gap of 100 bp
31467 85408: contig of 53942 bp in length
85409 85508: gap of 100 bp
85509 89621: contig of 4113 bp in length
89622 89721: gap of 100 bp
89722 104564: contig of 14843 bp in length
104565 104664: gap of 100 bp
104665 112440: contig of 7776 bp in length
112441 112540: gap of 100 bp
112541 131698: contig of 19158 bp in length
131699 131798: gap of 100 bp
131799 142284: contig of 10486 bp in length
142285 142384: gap of 100 bp
142385 151005: contig of 8621 bp in length
151006 151106: gap of 100 bp
151106 157224: contig of 6119 bp in length
157225 157324: gap of 100 bp

```

misc_feature      /estimated_length=100
166997..222799
/note="assembly_fragment"
gap              222800..222899
/estimated_length=100
misc_feature      222900..229569
/note="assembly_fragment"
gap              229570..229669

Query Match      54.1%; Score 27.6; DB 14; Length 329223;
Best Local Similarity 72.0%; Pred. No. 6.4;
Matches 36; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ACAGACTGCGAAGGAGGAGGCTCACTGAGGGCTCTGTCAACCCAGG 50
|||||
Db 199292 ACAGTCCGCGAGAGGCTGGAGGTCACAGAGTCCCTTTTCACACAGG 199341
|||||

RESULT 10
AC132271      173029 bp      DNA      linear      ROD 08-MAR-2005
LOCUS      Mus musculus BAC clone RP24-338E23 from 5, complete sequence.
DEFINITION      AC132271
ACCESSION      AC132271
VERSION      AC132271.5 GI:60301731
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 173029)
Tomlinson, C., Haglund, K. and Bielicki, L.
The sequence of Mus musculus BAC clone RP24-338E23
Unpublished (2001)
REFERENCE      2 (bases 1 to 173029)
McPherson, J. D. and Waterston, R. H.
Direct Submission
Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 173029)
Wilson, R. K.
Direct Submission
Submitted (09-FEB-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 173029)
Wilson, R. K.
Direct Submission
Submitted (26-FEB-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 173029)
Wilson, R. K.
Direct Submission
Submitted (08-MAR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Feb 26, 2005 this sequence version replaced gi:58801841.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
-----
Center project name: M_BB0338E23

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-24 BAC library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57Bl/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES
source
1..173029
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone_lib="RPCI-24"
/clone="RP24-338E23"
24291..24314
/note="Sequence derived from PCR product of project DNA."
misc_feature
24956..25048
/note="Sequence derived from one plasmid subclone."
unsure
45800..45813
/note="Sequence derived from one plasmid subclone."
misc_feature
139522..139794
/note="Sequence derived from PCR product of genomic DNA."
misc_feature
139840..140750
/note="Sequence derived from PCR product of genomic DNA."
unsure
140751..141175
/note="Unresolved simple sequence repeat."
misc_feature
141176..141456
/note="Sequence derived from PCR product of genomic DNA."
unsure
163787..163801
/note="Sequence derived from one plasmid subclone."

ORIGIN
Query Match      53.7%; Score 27.4; DB 9; Length 173029;
Best Local Similarity 75.6%; Pred. No. 7.5;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 CTGCCAAAGAGCAAGAGTCACTGAGGCTCTGTCAACCCAGGA 51
|||||
Db 167276 CTGCCAAAGAGCAAGTCACTGAGAGCTCTGCTCAAGGA 167320
|||||

RESULT 11
AC131679      198890 bp      DNA      linear      ROD 06-FEB-2004
LOCUS      Mus musculus BAC clone RP23-339I6 from chromosome 5, complete
sequence.
DEFINITION      AC131679
ACCESSION      AC131679.2 GI:40949631
VERSION      AC131679
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 198890)
Haakenson, W., Haglund, K., Bielicki, L. and Meyer, R.
The sequence of Mus musculus BAC clone RP23-339I6
Unpublished (2001)
REFERENCE      2 (bases 1 to 198890)
Wilson, R.
Sequencing of Mus musculus
Unpublished (2001)
JOURNAL

```


REFERENCE 3 (bases 1 to 198890)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 198890)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 198890)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (16-JAN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 198890)
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2004) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Jan 16, 2004 this sequence version replaced gi:22475611.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics

Center project name: M_BA0339106

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-23 BAC Library has been constructed by Kazutoyo Oregawa and Minako Tatenno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES
source
1..198890
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-33916"
/clone_lib="RPCI-23"
47..120
/rpt_family="ID"
297..460
repeat_region
repeat_region

repeat_region
1033..1208
/rpt_family="B4"
repeat_region
1462..1571
/rpt_family="MaLR"
repeat_region
2980..3174
/rpt_family="MIR"
repeat_region
3779..3963
/rpt_family="L1"
repeat_region
4280..5114
/rpt_family="B2"
repeat_region
5301..5581
/rpt_family="L1"
repeat_region
5886..6053
/rpt_family="B4"
repeat_region
6168..6368
/rpt_family="MaLR"
repeat_region
7206..7332
/rpt_family="B2"
repeat_region
7999..8016
/rpt_family="Alu"
8227..8357
/notes="Sequence derived from one plasmid subclone."
repeat_region
8720..8930
/rpt_family="B4"
repeat_region
8984..9066
/rpt_family="B4"
repeat_region
9912..10042
/rpt_family="B4"
repeat_region
10327..10743
/rpt_family="ERVK"
repeat_region
11314..11597
/rpt_family="B4"
repeat_region
11804..11912
/rpt_family="L1"
repeat_region
11930..12128
/rpt_family="RMER15"
repeat_region
12129..12320
/rpt_family="B2"
repeat_region
12321..12677
/rpt_family="RMER15"
repeat_region
12692..12840
/rpt_family="L1"
repeat_region
12990..13061
/rpt_family="MaLR"
repeat_region
13062..13245
/rpt_family="B2"
repeat_region
13246..13573
/rpt_family="MaLR"
repeat_region
13636..13829
/rpt_family="B2"
repeat_region
15079..15326
/rpt_family="MaLR"
repeat_region
16411..16566
/rpt_family="B4"
repeat_region
17194..17300
/rpt_family="Alu"
repeat_region
17404..17610
/rpt_family="MER1_type"
repeat_region
18519..18954
/rpt_family="MaLR"
repeat_region
21064..21238
/rpt_family="B2"
repeat_region
21800..21952
/rpt_family="Alu"
repeat_region
22082..22205
/rpt_family="Alu"
repeat_region
23272..23353
/rpt_family="Alu"
repeat_region
23446..23514
/rpt_family="Alu"
repeat_region
23665..23827
/rpt_family="L2"

```

repeat_region 25366..25417
/rpt_family="MER1_type"
repeat_region 25459..25611
/rpt_family="Alu"
repeat_region 27083..27314
/rpt_family="L1"
repeat_region 28449..28648
/rpt_family="B2"
repeat_region 28678..28866
/rpt_family="L1"
repeat_region 29380..29467
/rpt_family="Alu"
repeat_region 29938..30007
/rpt_family="ID"
repeat_region 30078..30353
/rpt_family="B4"
repeat_region 31735..31881
/rpt_family="B4"
repeat_region 31927..32015
/rpt_family="B4"
repeat_region 32016..32220
/rpt_family="B2"
repeat_region 32221..32322
/rpt_family="B4"
repeat_region 34803..34851
/rpt_family="Alu"
repeat_region 35801..35859
/rpt_family="ERV1"
repeat_region 36514..36689
/rpt_family="MER1_type"
repeat_region 36727..36998
/rpt_family="B4"
repeat_region 38298..38494
/rpt_family="ERVK"
repeat_region 38675..38730

Query Match 53.7%; Score 27.4; DB 9; Length 198890;
Best Local Similarity 75.6%; Pred. No. 7 5;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 CTGGCAAGAGGCAAGAGTCACTGAGGGCTCTGTCAACCCAGGA 51
|||||
Db 11491 CTGCCAAAAGCAACAGCTCACTGAGAGCCTCTGCTCAAGGA 11535

RESULT 12
AC094305/c
LOCUS AC094305 231980 bp DNA linear HTG 09-MAY-2003
DEFINITION Rattus norvegicus clone CH230-3P23, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC094305
AC094305.11 GI:30467414
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 231980)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,C., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Deigado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jarkson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kwis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuewa,L., Loulseghe,H., Iozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nair,L.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,R., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 231980)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 231980)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24952938.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAI2
Center clone name: CH230-3P23
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 205998 bases at least Q40

```

Consensus quality: 209782 bases at least Q30	
Consensus quality: 212122 bases at least Q20	
Estimated insert size: 220755; sum-of-contigs estimation	
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation	

* NOTE: Estimated insert size may differ from sequence length	
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).	
* NOTE: This is a 'working draft' sequence. It currently	
consists of 3 contigs. The true order of the pieces	
is not known and their order in this sequence record is	
arbitrary. Gaps between the contigs are represented as	
runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
be preserved.	
* 1 227830: contig of 227830 bp in length	
* 227831 227930: gap of unknown length	
* 227931 229095: contig of 1165 bp in length	
* 229096 229195: gap of unknown length	
* 229196 231980: contig of 2785 bp in length.	
Location/Qualifiers	
1. 231980	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clone="CH230-3F23"	
1. 1374	
/note="wgs end_extension	
clone_end:T7"	
91004. 91190	
/note="clone_boundary	
clone_end:T7	
site:ECORI	
end_sequence:BH309836"	
95949. 96064	
/note="clone_boundary	
clone_end:Sp6	
site:ECORI	
end_sequence:BH309839"	
226368. 227830	
/note="wgs end_extension	
clone_end:Sp6"	
227831. 227930	
/estimated_length=unknown	
229096. 229195	
/estimated_length=unknown	
ORIGIN	
Query Match 52.2%; Score 26.6; DB 14; Length 231980;	
Best Local Similarity 71.4%; Pred. NO. 16;	
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;	
QY 1 ACAGACTGGCAAGGCGACAGGTCCTACTGAGGGCCCTCTGCACCCAG 49	
Db 121661 ACACGATTGACAAAGTAGTAAGAGGGCACTGGCGCCACCTCTGCAGCAG 121613	
RESULT 13	
AC106528/c	
LOCUS	
DEFINITION	
6 unordered pieces.	
AC106528	
AC106528.4 GI:30580732	
HTG; HTGS_PHASE1; HTGS_DRAPT; HTGS_ENRICHED.	
SOURCE	
Rattus norvegicus (Norway rat)	
ORGANISM	
Rattus norvegicus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;	
Sciurognathi; Muridea; Muridae; Murinae; Rattus.	
1 (bases 1 to 261198)	
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,	
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,	
REFERENCE	
AUTHORS	
Consensus quality: 209782 bases at least Q30	
Consensus quality: 212122 bases at least Q20	
Estimated insert size: 220755; sum-of-contigs estimation	
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation	

* NOTE: Estimated insert size may differ from sequence length	
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).	
* NOTE: This is a 'working draft' sequence. It currently	
consists of 3 contigs. The true order of the pieces	
is not known and their order in this sequence record is	
arbitrary. Gaps between the contigs are represented as	
runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
be preserved.	
* 1 227830: contig of 227830 bp in length	
* 227831 227930: gap of unknown length	
* 227931 229095: contig of 1165 bp in length	
* 229096 229195: gap of unknown length	
* 229196 231980: contig of 2785 bp in length.	
Location/Qualifiers	
1. 231980	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clone="CH230-3F23"	
1. 1374	
/note="wgs end_extension	
clone_end:T7"	
91004. 91190	
/note="clone_boundary	
clone_end:T7	
site:ECORI	
end_sequence:BH309836"	
95949. 96064	
/note="clone_boundary	
clone_end:Sp6	
site:ECORI	
end_sequence:BH309839"	
226368. 227830	
/note="wgs end_extension	
clone_end:Sp6"	
227831. 227930	
/estimated_length=unknown	
229096. 229195	
/estimated_length=unknown	
ORIGIN	
Query Match 52.2%; Score 26.6; DB 14; Length 231980;	
Best Local Similarity 71.4%; Pred. NO. 16;	
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;	
QY 1 ACAGACTGGCAAGGCGACAGGTCCTACTGAGGGCCCTCTGCACCCAG 49	
Db 121661 ACACGATTGACAAAGTAGTAAGAGGGCACTGGCGCCACCTCTGCAGCAG 121613	
REFERENCE	
AUTHORS	
Consensus quality: 209782 bases at least Q30	
Consensus quality: 212122 bases at least Q20	
Estimated insert size: 220755; sum-of-contigs estimation	
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation	

* NOTE: Estimated insert size may differ from sequence length	
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).	
* NOTE: This is a 'working draft' sequence. It currently	
consists of 3 contigs. The true order of the pieces	
is not known and their order in this sequence record is	
arbitrary. Gaps between the contigs are represented as	
runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
be preserved.	
* 1 227830: contig of 227830 bp in length	
* 227831 227930: gap of unknown length	
* 227931 229095: contig of 1165 bp in length	
* 229096 229195: gap of unknown length	
* 229196 231980: contig of 2785 bp in length.	
Location/Qualifiers	
1. 231980	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	

table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLCS
Center clone name: CH230-4087
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 249514 bases at least Q40
Consensus quality: 251911 bases at least Q30
Consensus quality: 253574 bases at least Q20
Estimated insert size: 269546; sum-of-contrigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contrigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 5789: contig of 5789 bp in length
* 5790 5899: gap of unknown length
* 255153: contig of 249264 bp in length
* 255154 255253: gap of unknown length
* 255254 256303: contig of 1050 bp in length
* 256304 256403: gap of unknown length
* 256404 257841: contig of 1438 bp in length
* 257842 257941: gap of unknown length
* 257942 259377: contig of 1436 bp in length
* 259378 259477: gap of unknown length
* 259478 261198: contig of 1721 bp in length.
*
* Location/Qualifiers
* 1. .261198
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-4087"
* 1. .3720
* /notes="wgs contig"
* /notes="wgs contig"
* 3771. .5789
* /notes="wgs contig"
* 5790. .5899
* /estimated_length=unknown
* 5890. .6946
* /notes="wgs contig"
* complement(243387. .244025)
* /notes="clone boundary
* clone end:T7"
* site:EcoRI
* end_sequence:BH315060"
* 252740. .255153
* /notes="wgs end_extension
* clone end:T7"
* 255154. .255253
* /estimated_length=unknown
* 256304. .256403
* /estimated_length=unknown
* 257842. .257941
* /estimated_length=unknown
* 259378. .259477
* /estimated_length=unknown
*
* ORIGIN
Query Match 52.2%; Score 26.6; DB 14; Length 261198;
Best Local Similarity 71.4%; Pred. No. 16;

```

```

Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGCGCTCTGTCAACCCAG 49
Db 183141 ACACGATTGACAAAGATAAGAGGCGCACTGCGCCACCTGTGCAGCAG 183093

RESULT 14
AC149647 282320 bp DNA linear HTG 01-JUL-2005
Bos taurus clone CH240-68K2, WORKING DRAFT SEQUENCE, 22 unordered
pieces.
AC149647.3 GI:68302937
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Bos taurus (cow)
Bos taurus
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 282320)
Muzny,D.Marie., Metaker,M.Lee., Abranzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,W., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensunewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidas,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokeme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puafo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willison,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,J., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 282320)
TITLE

```

AUTHORS
TITLE
JOURNAL

Worley, K.C.
Direct Submission
Submitted (19-JUN-2004) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 282320)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

On Jun 29, 2005 this sequence version replaced gi:50540557.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FDTZ
Center clone name: CH240-68K2
----- Summary Statistics
Assembly program: Atlas 3.0:
Consensus quality: 23336 bases at least Q40
Consensus quality: 235130 bases at least Q30
Consensus quality: 236805 bases at least Q20
Estimated insert size: 239331; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 22 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence.
as soon as it is available and the accession number will
be preserved.

1
7906: contig of 7906 bp in length
7907
7956: gap of 50 bp
7957
21416: contig of 13460 bp in length
21417
21466: gap of 50 bp
32425: contig of 10959 bp in length
32426
32475: gap of 50 bp
32476
71121: contig of 38646 bp in length
71122
71171: gap of 50 bp
73347: contig of 2176 bp in length
73348
107554: gap of 34207 bp
123340: contig of 15786 bp in length
123341
123390: gap of 50 bp
127911: contig of 4521 bp in length
127912
127961: gap of 50 bp
127962
135626: contig of 7665 bp in length
135627
135677: gap of 50 bp
144612: contig of 8936 bp in length
144663
144662: gap of 50 bp
145739: contig of 1077 bp in length
145740
151919: gap of 6180 bp
158564: contig of 6645 bp in length
158565
158614: gap of 50 bp
158615
182943: contig of 24329 bp in length

182944
183044
184051: contig of 1008 bp in length
184052
184151: gap of unknown length
184152
185427: contig of 1276 bp in length
185428
185527: gap of unknown length
185528
185527: contig of 1069 bp in length
185597
186696: gap of unknown length
186697
188035: contig of 1339 bp in length
188036
188135: gap of unknown length
188136
189300: contig of 1165 bp in length
189301
189400: gap of unknown length
189401
191191: contig of 1791 bp in length
191192
191291: gap of unknown length
191292
198350: contig of 6959 bp in length
198351
198350: gap of unknown length
198351
203718: contig of 5368 bp in length
203719
203818: gap of unknown length
203819
209340: contig of 5522 bp in length
209341
209440: gap of unknown length
209441
282320: contig of 72880 bp in length.

FEATURES

source

1. 282320
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-68K2"
7907. 7956
/estimated_length=50
21417. 21466
/estimated_length=50
32426. 32475
/estimated_length=50
71122. 71171
/estimated_length=50
73348. 107554
/estimated_length=34207
123341. 123390
/estimated_length=50
127912. 127961
/estimated_length=50
135627. 135676
/estimated_length=50
144613. 144662
/estimated_length=50
145740. 151919
/estimated_length=6180
158565. 158614
/estimated_length=50
182944. 183043
/estimated_length=unknown
184052. 184151
/estimated_length=unknown
185428. 185527
/estimated_length=unknown
186597. 186696
/estimated_length=unknown
188036. 188135
/estimated_length=unknown
189301. 189400
/estimated_length=unknown
191192. 191291
/estimated_length=unknown
198251. 198350

Query Match 52.2%; Score 26.6; DB 14; Length 282320;

Best Local Similarity 71.4%; Pred. No. 16;

Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACGTGAGGGCCCTCTGTCCACCAGG 50

DB 45334 CAGCACTGGCAAGAGGCAAGAGGTCACGTGAGGGCCCTCTGTCCACCAGG 45382

```

RESULT 15
AC008459      136581 bp      DNA      linear      PRI 04-FEB-2003
LOCUS      Homo sapiens chromosome 5 clone CTC-352D11, complete sequence.
DEFINITION      AC008459
ACCESSION      AC008459
VERSION        AC008459.7 GI:28201472
KEYWORDS      HTG.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE      1 (bases 1 to 136581)
              DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS      DOE Joint Genome Institute
TITLE        Direct Submission
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 136581)
              DOE Joint Genome Institute.
AUTHORS      DOE Joint Genome Institute
TITLE        Direct Submission
JOURNAL      Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE      3 (bases 1 to 136581)
              DOE Joint Genome Institute.
AUTHORS      DOE Joint Genome Institute
TITLE        Direct Submission
JOURNAL      Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE      4 (bases 1 to 136581)
              DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS      DOE Joint Genome Institute
TITLE        Direct Submission
JOURNAL      Submitted (04-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
              Drive, Walnut Creek, CA 94598, USA
COMMENT      On Feb 4, 2003 this sequence version replaced gi:14579681.
              Draft Sequence Produced by DOE Joint Genome Institute
              www.jgi.doe.gov
              Finishing Completed at Stanford Human Genome Center
              www-shgc.stanford.edu
              Quality: Phrap Quality >=40 99.9% of Sequence;
              Estimated Total Number of Errors is 0.1.
              NOTE: Shatter libraries failed to verify the dinucleotide repeat
              region 3558-5794. Unsure number of repeat copies 3558-5794.
              Forced join at 4399.

FEATURES             Location/Qualifiers
     source            1..136581
                       /organism="Homo sapiens"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:9606"
                       /chromosome="5"
                       /clone="CTC-352D11"
     misc_feature      3558..5794
                       /note="NOTE: Shatter libraries failed to verify the
                       dinucleotide repeat region 3558-5794. Unsure number of
                       repeat copies 3558-5794. Forced join at 4399."

ORIGIN
Query Match      51.4%; Score 26.2; DB 8; Length 136581;
Best Local Similarity 79.5%; Pred. No. 23;
Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      7      CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTGAC 45
          ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      19816 CTGCAGATAAGCAGGAGGCCCTCTGGAGGCCTCTGTGAC 19854

```

Search completed: February 7, 2006, 15:20:44
 Job time : 281.097 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:30:54 ; Search time 10.8591 Seconds
(without alignments)
8348.349 Million cell updates/sec

Title: US-10-623-108-5_COPY_1186_1236

Perfect score: 51

Sequence: 1 acgaactggcaagaggca.....aggcctctgtcaccaggga 51

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H/COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25.6	50.2	99916	3	US-09-816-095-3
2	24.4	47.8	601	3	US-09-949-016-122934
3	24.4	47.8	601	3	US-09-949-016-123028
4	24.4	47.8	601	3	US-09-949-016-123122
5	24.4	47.8	601	3	US-09-949-016-123216
6	24.4	47.8	601	3	US-09-949-016-123310
7	24.4	47.8	1023	3	US-09-902-540-9591
8	24.4	47.8	13346	3	US-09-902-540-1089
9	24.4	47.8	32379	3	US-09-949-016-15217
10	24.4	47.8	32379	3	US-09-949-016-15218
11	24.4	47.8	32379	3	US-09-949-016-15219
12	24.4	47.8	32379	3	US-09-949-016-15220
13	24.4	47.8	32379	3	US-09-949-016-15221
14	24.4	47.8	32379	3	US-09-949-016-15222
15	24.4	47.8	32379	3	US-09-949-016-15223
16	24.4	47.8	32379	3	US-09-949-016-15224
17	24.4	47.8	32379	3	US-09-949-016-15225
18	24.4	47.8	32379	3	US-09-949-016-15226
19	24.4	47.8	64377	3	US-09-949-016-15212
20	24.4	47.8	64377	3	US-09-949-016-15213
21	24.4	47.8	64377	3	US-09-949-016-15214
22	24.4	47.8	64377	3	US-09-949-016-15215
23	24.4	47.8	64377	3	US-09-949-016-15216
24	23.2	45.5	3089	3	US-10-104-047-914

25	22.8	44.7	1191	3	US-09-902-540-8380	Sequence 8380, Ap
c	22.8	44.7	2022	3	US-09-023-655-416	Sequence 416, App
c	22.8	44.7	2088	3	US-09-949-016-4237	Sequence 4237, Ap
28	22.8	44.7	2737	3	US-09-902-540-3798	Sequence 3798, App
c	22.8	44.7	4231	3	US-09-949-016-809	Sequence 809, App
c	22.8	44.7	7035	3	US-09-902-540-878	Sequence 878, App
30	22.8	44.7	18551	3	US-09-902-540-1187	Sequence 1187, Ap
31	22.8	44.7	21593	3	US-09-949-016-15979	Sequence 15979, A
c	22.8	44.7	96878	3	US-09-949-016-12551	Sequence 12551, A
34	22.8	44.7	114793	3	US-10-148-806-3	Sequence 187024, A
35	22.6	44.3	601	3	US-09-949-016-187024	Sequence 187025, A
36	22.6	44.3	601	3	US-09-949-016-187025	Sequence 187026, A
37	22.6	44.3	601	3	US-09-949-016-187026	Sequence 187026, A
c	22.6	44.3	56616	3	US-09-949-016-12462	Sequence 12462, A
c	22.6	44.3	56616	3	US-09-949-016-17085	Sequence 17085, A
39	22.6	44.3	152070	3	US-09-949-016-15402	Sequence 15402, A
c	22.6	44.3	601	3	US-09-949-016-120678	Sequence 120678, A
41	22.4	43.9	601	3	US-09-949-016-158390	Sequence 158390, A
42	22.4	43.9	601	3	US-09-949-016-158391	Sequence 158391, A
43	22.4	43.9	601	3	US-09-949-016-158392	Sequence 158392, A
44	22.4	43.9	601	3	US-09-949-016-158392	Sequence 158392, A
45	22.4	43.9	601	3	US-09-949-016-158393	Sequence 158393, A

ALIGNMENTS

RESULT 1

US-09-816-095-3
; Sequence 3, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816, 095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(99916)
; OTHER INFORMATION: n = A, T, C or G
US-09-816-095-3

Query Match	50.2%	Score	25.6;	DB	3;	Length	99916;
Best Local Similarity	70.8%	Pred. No.	10;				
Matches	34;	Conservative	0;	Mismatches	14;	Indels	0;
Gaps	0;						
Qy	2	CAGACTGGCAAGAGCAAGAGTCTACTAGGGCTCTCTCACCAG	49				
Db	1819	CAGACCGCGGAGAGCAAGAGGCTCTCTCACCAG	1866				

RESULT 2

US-09-949-016-122934
; Sequence 122934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-122934

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 3
US-09-949-016-123028
; Sequence 123028, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123028
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123028

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 4
US-09-949-016-123122
; Sequence 123122, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123122
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123122

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 5
US-09-949-016-123216
; Sequence 123216, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123216
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123216

Query Match          47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTC 43
Db 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGGAATCTATC 177

RESULT 6
US-09-949-016-123310
; Sequence 123310, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123310
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-123310
```


Query Match 47.8%; Score 24.4; DB 3; Length 601;
Best Local Similarity 73.8%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGAGGCTCTGTC 43
DB 136 CAGGACTGACACAGGGGCTACAGGCCACTGTGGAAATCTATC 177

RESULT 7

US-09-902-540-9591/c
; Sequence 9591, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9591
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-9591

Query Match 47.8%; Score 24.4; DB 3; Length 1023;
Best Local Similarity 68.0%; Pred. No. 11;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGAGGCTCTGTCACCCAGGA 51
DB 866 CAGAACTGGCGGCGCCAGCAGCAGTCACTCAGGGGCTCCAACTCCAGGA 817

RESULT 8

US-09-902-540-1089/c
; Sequence 1089, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1089
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1089

Query Match 47.8%; Score 24.4; DB 3; Length 1346;
Best Local Similarity 68.0%; Pred. No. 19;
Matches 34; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGAGGCTCTGTCACCCAGGA 51
DB 4227 CAGAACTGGCGGCGCCAGCAGCAGTCACTCAGGGGCTCCAACTCCAGGA 4178

RESULT 9

US-09-949-016-15217/c
; Sequence 15217, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15217
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15217

Query Match 47.8%; Score 24.4; DB 3; Length 32379;
Best Local Similarity 73.8%; Pred. No. 23;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGAGGCTCTGTC 43
DB 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGAAATCTATC 1289

RESULT 10

US-09-949-016-15218/c
; Sequence 15218, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15218
; LENGTH: 32379
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15218

Query Match 47.8%; Score 24.4; DB 3; Length 32379;
Best Local Similarity 73.8%; Pred. No. 23;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGAACTGGCAAGAGGCAAGAGTCACTAGAGGCTCTGTC 43
DB 1330 CAGGACTGACACAGGGGCTACAGGCCACTGTGGAAATCTATC 1289

RESULT 11

US-09-949-016-15219/c
; Sequence 15219, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

This Page Blank (uspto)


```
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SEQ ID NO 34830
; TYPE: DNA
; LENGTH: 1656
; ORGANISM: Bovine 19866880752334
US-10-750-623-34830

Query Match      47.5%; Score 24.2; DB 7; Length 1656;
Best Local Similarity 71.1%; Pred. No. 7.9;
Matches 32; Conservative 0; Mismatches 0; Indels 13; Gaps 0;

QY 5 AACTGGCAAGGCGAAGGTCACGTGAGGGCCTCTGTCAACCAG 49
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 578 AACAGTCAGTGAGATAGATCGAAGAGAGCCTCTGCCACCCAG 622

RESULT 3
US-10-750-185-20700
; Sequence 20700, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20700
; TYPE: DNA
; LENGTH: 600
; ORGANISM: Bovine MMBT20675
US-10-750-185-20700

Query Match      46.3%; Score 23.6; DB 7; Length 600;
Best Local Similarity 76.3%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 9; Indels 9; Gaps 0;

QY 1 ACAGAACTGGCAAGGCGAAGGTCACGTGAGGGCCT 38
    ||||| ||||| ||||| ||||| ||||| |||||
Db 559 ATAAACCTGGCAAGGCGAAGGAGGACCCCTGAGGTCT 596

RESULT 4
US-10-750-623-20700
; Sequence 20700, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
```

```
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20700
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT20675
US-10-750-623-20700

Query Match      46.3%; Score 23.6; DB 7; Length 600;
Best Local Similarity 76.3%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 9; Indels 9; Gaps 0;

QY 1 ACAGAACTGGCAAGGCGAAGGTCACGTGAGGGCCT 38
    ||||| ||||| ||||| ||||| ||||| |||||
Db 559 ATAAACCTGGCAAGGCGAAGGAGGACCCCTGAGGTCT 596

RESULT 5
US-10-750-185-40559
; Sequence 40559, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40559
; LENGTH: 903
; TYPE: DNA
; ORGANISM: Bovine 19866881097008
US-10-750-185-40559

Query Match      46.3%; Score 23.6; DB 7; Length 903;
Best Local Similarity 76.3%; Pred. No. 12;
Matches 29; Conservative 0; Mismatches 9; Indels 9; Gaps 0;

QY 1 ACAGAACTGGCAAGGCGAAGGTCACGTGAGGGCCT 38
    ||||| ||||| ||||| ||||| ||||| |||||
Db 780 ATAAACCTGGCAAGGCGAAGGAGGACCCCTGAGGTCT 817

RESULT 6
US-10-750-623-40559
; Sequence 40559, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
```

```

; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40559
; LENGTH: 903
; TYPE: DNA
; ORGANISM: Bovine 19866881097008
US-10-750-623-40559

Query Match      46.3%; Score 23.6; DB 7; Length 903;
Best Local Similarity 76.3%; Pred. No. 12; Mismatches 0; Gaps 0;
Matches 29; Conservative 0; Indels 9;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCT 38
    |||||
DB 780 ATAACTGGCAAGAGGGAAGGAGCCCTGAGGTCT 817

RESULT 7
US-11-124-368A-3178/c
; Sequence 3178, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: C0001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3178
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-3178

Query Match      45.9%; Score 23.4; DB 8; Length 201;
Best Local Similarity 64.7%; Pred. No. 10; Mismatches 1; Gaps 0;
Matches 33; Conservative 1; Indels 17;

QY 1 ACAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGGA 51
    |||||
DB 120 AAACAATTTTCAGGGAGTCRTGAAGCCCTTCTGGGCCCTCTTCCACCCAGGA 70

RESULT 8
US-10-775-169-316/c
; Sequence 316, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 316
; LENGTH: 34875
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-775-169-316

Query Match      45.5%; Score 23.2; DB 7; Length 34875;
Best Local Similarity 70.5%; Pred. No. 35; Mismatches 13; Indels 0; Gaps 0;
Matches 31; Conservative 0;

QY 7 CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGG 50
    |||||
DB 12441 CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGG 12398

RESULT 9
US-11-121-086-4/c
; Sequence 4, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 164810
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-4

Query Match      45.5%; Score 23.2; DB 8; Length 164810;
Best Local Similarity 70.5%; Pred. No. 48; Mismatches 13; Indels 0; Gaps 0;
Matches 31; Conservative 0;

QY 7 CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGG 50
    |||||
DB 8707 CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGG 8664

RESULT 10
US-11-121-086-6/c
; Sequence 6, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 172543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-6

Query Match      45.5%; Score 23.2; DB 8; Length 172543;
Best Local Similarity 70.5%; Pred. No. 48; Mismatches 13; Indels 0; Gaps 0;
Matches 31; Conservative 0;

QY 7 CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGG 50
    |||||
DB 143325 CTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGG 143282

RESULT 11
US-10-750-185-64443/c
```



```

; LENGTH: 182303
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-45

Query Match      44.7%; Score 22.9; DB 8; Length 182303;
Best Local Similarity 66.0%; Pred. No. 68;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy      2 CAGAACTGGCAAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCACCACAGGA 51
      ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 145227 CAGAACATGGAGGAAGTGAAGAGGAGGGGCTTCGGGTGCCCATGA 145178

```

Search completed: February 7, 2006, 14:11:06
Job time : 22.0694 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:27:59 ; Search time 9908.9 Seconds
(without alignments)
10538.142 Million cell updates/sec

US-10-623-108-5

Title: 1837

Perfect score: 1

Sequence: 1 gagagtgagatggcggcg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_to.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1725	93.9	1813	8	BC013992		BC013992 Homo sapi
2	1708.4	93.0	1866	6	AR164078		AR164078 Sequence
3	1708.4	93.0	1866	6	BD058089		BD058089 Antisense
4	1708.4	93.0	1866	6	CS063502		CS063502 Sequence
5	1708.4	93.0	1866	6	CS081283		CS081283 Sequence
6	1708.4	93.0	1866	6	AX770530		AX770530 Sequence
7	1708.4	93.0	1866	6	HSERK1		X60188 Human ERK1
8	1708.4	93.0	3732	6	AX676347		AX676347 Sequence
9	1632	88.8	1745	8	HUMKER1A		M84490 Human extra
10	1546.8	84.2	2567	8	AK091009		AK091009 Homo sapi
11	1540.8	83.9	2722	8	HSM806004		BX537897 Homo sapi
12	1201	65.4	1284	8	HS44KDAP		Z11696 H.sapiens 4
13	1140	62.1	1140	11	AY888192		AY888192 Synthetic
14	1140	62.1	1140	11	AY890458		AY890458 Synthetic
15	1140	62.1	1140	11	AY893270		AY893270 Synthetic
16	1138.4	62.0	1140	6	AX722274		AX722274 Synthetic
17	1138.4	62.0	1140	11	AY335595		AY335595 Synthetic
18	1138.4	62.0	1140	11	AY890764		AY890764 Synthetic

19	1138.4	62.0	1140	11	AY890765		AY890765 Synthetic
20	1138.4	62.0	1896	6	A84423		A84423 Sequence 38
21	1138.4	62.0	1896	6	BD082524		BD082524 A method
22	1138.4	62.0	1896	6	BD228337		BD228337 An improv
23	1138.4	62.0	1896	6	CQ829947		CQ829947 Sequence
24	1138.4	62.0	1896	6	AR279902		AR279902 Sequence
25	1138.4	62.0	1896	6	AX427322		AX427322 Sequence
26	1080.2	58.8	1125	6	AX803349		AX803349 Sequence
27	1061	57.8	1783	9	SC029712		SC029712 Mus muscu
28	1060.4	57.7	1875	9	S46779		S46779 extracellul
29	1049.6	57.1	1736	9	RATERK1A		M61177 Rattus norv
30	1048	57.0	1721	9	RNU12008		Z14249 M.musculus
31	1033.4	56.3	1718	9	MMERK1		M38194 Rat extrace
32	1030.2	56.1	1715	9	RATERK1		AR016531 Sequence
33	1025	55.8	1747	6	AR016531		AR036244 Sequence
34	1025	55.8	1747	6	AR036244		AR171290 Sequence
35	1025	55.8	1747	6	AR171290		AR172791 Sequence
36	1025	55.8	1747	6	AR172791		I34159 Sequence 1
37	1025	55.8	1747	6	I34199		AY033607 Homo sapi
38	1015.4	55.3	1074	8	AY033607		X65198 R.norvegicu
39	965.4	52.6	1238	9	RNM4PKIN		A84441 Sequence 56
40	958.6	52.2	1875	6	A84441		BD082533 A method
41	958.6	52.2	1875	6	BD082533		CQ829965 Sequence
42	958.6	52.2	1875	6	CQ829965		AR279911 Sequence
43	958.6	52.2	1875	6	AR279911		AX427340 Sequence
44	958.6	52.2	1875	6	AX427340		BC013754 Mus muscu
45	939.6	51.1	1632	9	BC013754		

ALIGNMENTS

RESULT 1	BC013992	1813 bp	linear	PRI 29-JUN-2004
LOCUS	Homo sapiens mitogen-activated protein kinase 3, mRNA (cdna clone			
DEFINITION	MGC:20180 IMAGE:3634492), complete cds.			
ACCESSION	BC013992			
VERSION	BC013992.1	GI:15559270		
KEYWORDS	MGC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 1813)			
AUTHORS	Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzyzinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
PUBMED	12477932			
REFERENCE	2 (bases 1 to 1813)			
AUTHORS	Strausberg, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (10-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer			

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter N., Ayèle, K., Beckstrom-Sternberg S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Skaterlipop, S., Thomas, P.J., Touchman, J.W.,
Tsongeon, C., Vogt, J.-L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A.C., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 28 Row: d Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES
SOURCE

```

Location/Qualifiers
1. .1813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20180.1"
/tissue_type="Lymphoid tissue"
/clone_lib="NIH MGC"
/lab_host="DH10B-R"
/note="Vector: pOTB"
1. .1813

```

gene

CDs

```

1 1813
/ gene="WAPK3"
/ note="synonyms: P44ERK1, P44MAPK, ERK1, PRK3"
/db_xref="GeneID:5595"
/db_xref="MIM:601795"
12. 1151
/ gene="WAPK3"
/ codon_start=1
/ product="mitogen-activated protein kinase 3"
/ protein_id="AAH13992.1"
/db_xref="GI:15559271"
/db_xref="GeneID:5595"
/db_xref="MIM:601795"
/ translation="MAAAAGCGGGERPRRTGCVGPVGEVMVKG
LQYIGEGAGWVSADYHVRKTRVAKIKSPHEQYTCORTLEIQI
IRDLRTSLBEMRDVYTVQDMLTDYLLKSLQSLNDHI CVFYLO
NVLKRDLPKNLLNTTCDLKI CDFGLARIADPDHHTGTFLTEYVA
SKGYTKSIDISVSGQILAEMLSNRIFPFKHYLDQNLHILGLGSPS
ARNYQLSPSTKTVKAWAKLPKSDSKALDLDMLTFNPKNKLTVEE
DPDPEYBESPEFAMELDDIIPKBLKELI IFOETARFQGVLEAP"

```

ORIGIN

	Query Match	93.9%;	Score 1725;	DB 8;	Length 1813;
	Best Local Similarity	97.2%;	Pred. No. 8.6e-293;		
	Matches 1786;	Conservative 0;	Mismatches 0;	Indels 51;	Gaps 1;
Qy	1	GAGGAGTGGAGATGCGCGCGCGCGGCTCAGCGGGGCGGGGGCGGGAGGCCCGGTAGAA	60		
Db	1	GAGGAGTGGAGATGCGCGCGCGCGGCTCAGCGGGGCGGGGGCGGGAGGCCCGGTAGAA	60		
Qy	61	CCGAGGGGGTCCGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTTCG	120		
Db	61	CCGAGGGGGTCCGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTTCG	120		
Ov	121	ACGTGGGCCCGCGCTACACGCAGTTGCAGTACATCTGGCGAGGGGCGCGTACGGCATGCTCA	180		

Db	121	ACGTGGGCCCGCGTTACCGCAGTTTGCAGTACATCTGGCGAGGGCGGTACGGCATGGTCA	180
Qy	181	GCTCGGCCTATGACCAACGTGCGCAAGACTCGCTGGCCATCAAGAAGATCAGCCCCCTTCG	240
Db	181	GCTCGGCCTATGACCAACGTGCGCAAGACTCGCTGGCCATCAAGAAGATCAGCCCCCTTCG	240
Qy	241	AAATCATCAGACTTACTGCCAGGCGACGCTCGGGAGATCCAGATCTCTGCTGGCTTCCGCC	300
Db	241	AAATCATCAGACTTACTGCCAGGCGACGCTCGGGAGATCCAGATCTCTGCTGGCTTCCGCC	300
Qy	301	ATGAGATGTCATCGGCATCCGAGACATCTCGGGCGTCCACCTCGGAAGCCATGAGAG	360
Db	301	ATGAGATGTCATCGGCATCCGAGACATCTCGGGCGTCCACCTCGGAAGCCATGAGAG	360
Qy	361	ATGTCTACATTTGTGCAGGACTGATGAGAGCTGACTGTGTAACAAGTTCTGTAAGAAGCCAGC	420
Db	361	ATGTCTACATTTGTGCAGGACTGATGAGAGCTGACTGTGTAACAAGTTCTGTAAGAAGCCAGC	420
Qy	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACACAGATCTCTGCGGGGCTCAAGTACA	480
Db	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACACAGATCTCTGCGGGGCTCAAGTACA	480
Qy	481	TCCACTCCGCCAAAGTGTCTCCACCGAGATCTAAAGCCCTCAACCTGCTCATCAACACCA	540
Db	481	TCCACTCCGCCAAAGTGTCTCCACCGAGATCTAAAGCCCTCAACCTGCTCATCAACACCA	540
Qy	541	CCTGCGACCTTAAGATTTGTGATTTCTGGCTGCGCCGGATTTCCGATCTCTGAGCATGACC	600
Db	541	CCTGCGACCTTAAGATTTGTGATTTCTGGCTGCGCCGGATTTCCGATCTCTGAGCATGACC	600
Qy	601	ACACCGGCTTCTTGACGAGGTATGTGGCTGCTACGGCTGTGTAACGGGCCCCAGAGATCATGC	660
Db	601	ACACCGGCTTCTTGACGAGGTATGTGGCTGCTACGGCTGTGTAACGGGCCCCAGAGATCATGC	660
Qy	661	TGAACTCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTCGTGGTG	720
Db	661	TGAACTCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTCGTGGTG	720
Qy	721	AGATGCTCTTAACCGGCCCATCTTCCCTCGCAAGCACTACCTGGATCAGCTCAACACCA	780
Db	721	AGATGCTCTTAACCGGCCCATCTTCCCTCGCAAGCACTACCTGGATCAGCTCAACACCA	780
Qy	781	TTCTGGGCATCTGGGCTCCCATCCACGAGGACCTGAATTTGATCATCAACATGAAGG	840
Db	781	TTCTGGGCATCTGGGCTCCCATCCACGAGGACCTGAATTTGATCATCAACATGAAGG	840
Qy	841	CCGGAATCTACCTACAGTCTCTGCGCTCACAAGAGGTGGCTTGGGCCAAGCTTTTCC	900
Db	841	CCGGAATCTACCTACAGTCTCTGCGCTCACAAGAGGTGGCTTGGGCCAAGCTTTTCC	900
Qy	901	CCAAGTCAGACTCCAAGCCCTTGACCTGTGACCCGATGTTAACTTTAAACCCCAATA	960
Db	901	CCAAGTCAGACTCCAAGCCCTTGACCTGTGACCCGATGTTAACTTTAAACCCCAATA	960
Qy	961	AACCGATCAAGTGGAGGAAGCGTGGCTCAACCTTACCTGGAGCAGTATATGACCCCGA	1020
Db	961	AACCGATCAAGTGGAGGAAGCGTGGCTCAACCTTACCTGGAGCAGTATATGACCCCGA	1020
Qy	1021	CGGATGACGAGTGGCGGAGGCCCTTACCTTCGCATGGAGCTGGATGACCTTACCTA	1080
Db	1021	CGGATGACGAGTGGCGGAGGCCCTTACCTTCGCATGGAGCTGGATGACCTTACCTA	1080
Qy	1081	AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGAGTGTGG	1140
Db	1081	AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGAGTGTGG	1140
Qy	1141	AGGCCCCCTTAGCCAGACAGACATCTTGCAACCTTGGGGCTGGAAACAGAACTGGCAAG	1200
Db	1141	AGGCCCCCTTAGCCAGACAGACATCTCTGCAACCTTGGGGCTGGAAACAGAACTGGCAAG	1180
Qy	1201	AGGCAAGAGGTCACTGAGGGGCTCTGTGATCCACGAGACTGCTGCTCTGCTGCCCTCTTCC	1260

Db	1204	GCCCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCCT-----	1243
Qy	1203	GCAGAGGTCACTGAGGGCTCTGTACCCAGGACCTGCTCTGCTGCCCTCTCCCG	1262
Db	1244	-----GGACCTGCTCTGCTGCTGCCCTCTCCCG	1272
Qy	1263	CCAGACTGTTAGAAATGGACACTGTGCCAGCCCGGACCTTGGCAGCCAGCCGGGGT	1322
Db	1273	CCAGACTGTTAGAAATGGACACTGTGCCAGCCCGGACCTTGGCAGCCAGCCGGGGT	1332
Qy	1323	GGAGCATGGGCTGGCCACTCTCTCTTTGTGAGGCTTCAGGCTTCAGGAGGCCAAG	1382
Db	1333	GGAGCATGGGCTGGCCACTCTCTCTTTGTGAGGCTTCAGGCTTCAGGAGGCCAAG	1392
Qy	1383	GCCTTCTCTCCCAACCCGCTCCCAAGGGCTCTGGGAGCTCAGGTGGCCCAAGTTC	1442
Db	1393	GCCTTCTCTCCCAACCCGCTCCCAAGGGCTCTGGGAGCTCAGGTGGCCCAAGTTC	1452
Qy	1443	AATCTCCC--GCTGTGCTGTGCGCCCTTACCTTCCCAGCGTCCCAAGTCTCTGGCAG	1499
Db	1453	AATCTCCCAGTGTGCTGTGCTGCGCCCTTACCTTCCCAGCGTCCCAAGTCTCTGGCAG	1512
Qy	1500	TTCTGGAATGGAGGGTTCTGGCTGCCCAACCTCTGTAAGGCGCAGAGTGGAGGTGG	1559
Db	1513	TTCTGGAATGGAGGGTTCTGGCTGCCCAACCTCTGTAAGGCGCAGAGTGGAGGTGG	1572
Qy	1560	GGGCGCTGAGTAGGAGACTCAGGCGCATGCTGCGCCCTCATCTCATTAACCCCAACC	1619
Db	1573	GGGCGCTGAGTAGGAGACTCAGGCGCATGCTGCGCCCTCATCTCATTAACCCCAACC	1632
Qy	1620	TAGTTTCCCTGAAGGAACATTCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGGC	1679
Db	1633	TAGTTTCCCTGAAGGAACATTCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGGC	1692
Qy	1680	GGGCCGAATCCCTCCCTGTCAAAGCTGTCACTTCGCGTCCCTCGCTCTCTGTGTGT	1739
Db	1693	GGGCCGAATCCCTCCCTGTCAAAGCTGTCACTTCGCGTCCCTCGCTCTCTGTGTGT	1752
Qy	1740	GGTGACAGAGTGGAGCTGGGGGCTGGAGAGCCCGGGCCCTGCCACCTCCCTGAC	1799
Db	1753	GGTGACAGAGTGGAGCTGGGGGCTGGAGAGCCCGGGCCCTGCCACCTCCCTGAC	1812
Qy	1800	CCGCTCAATATATAAATATAGATGTGTCTATGGCTG	1837
Db	1813	CCGCTCAATATATAAATATAGATGTGTCTATGGCTG	1850
RESULT 3			
BD058089			
LOCUS			
DEFINITION			
Antisense oligonucleotides for mitogen-activated protein kinases as			
therapy for breast cancer.			
ACCESSION			
BD058089			
VERSION			
BD058089.1 GI:22603695			
KEYWORDS			
JP 2001518881-A/1.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 1866)			
Sivaraman,V.S., Wang,H.Y. and Malbon,C.C.			
Antisense oligonucleotides for mitogen-activated protein kinases as			
therapy for breast cancer			
Patent: JP 2001518881-A 1 16-OCT-2001;			
THE RESEARCH FOUNDATION OF STATE UNIV OF NEW YORK			
COMMENT			
OS Homo sapiens (human)			
PN JP 2001518881-A/1			
PD 16-OCT-2001			
PF 19-MAR-1998 JP 1998541700			
PI VINALLA S SIVARAMAN,HSIEN YU WANG,CRAIG C MALBON PC			
C12N15/11,A61K31/70,C12Q1/68//A61K48/00			
CC The molecular type is cDNA.			

FEATURES		FH	Key	Location/Qualifiers	Location/Qualifiers
source				1..1866	
				/organism="Homo sapiens"	
				/mol_type="genomic DNA"	
				/db_xref="taxon:9606"	
ORIGIN					
Query Match			93.0%;	Score 1708.4;	DB 6; Length 1866;
Best Local Similarity			97.0%;	Pred. No. 7e-290;	
Matches 1783; Conservative			0;	Mismatches 1;	Indels 54; Gaps 2;
QY	3	GGAGTGGAGATGGCGCGCGCGCTCAGGGGGCGGGGCGGGGAGCCCGTGAAGACC	62		
DB	64	GGAGTGGAGATGGCGCGCGCGCTCAGGGGGCGGGGCGGGGAGCCCGTGAAGACC	123		
QY	63	GAGGGGGTGGCGCGCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGTTTCGAC	122		
DB	124	GAGGGGGTGGCGCGCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGTTTCGAC	183		
QY	123	GTGGGCGCGCGCTACACGCAGTTGCAGTACATCGGCGAGGGCGCGTACGSCATGGTCAGC	182		
DB	184	GTGGGCGCGCGCTACACGCAGTTGCAGTACATCGGCGAGGGCGCGTACGSCATGGTCAGC	243		
QY	183	TCGGCTATGACCACTGCGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCGAA	242		
DB	244	TCGGCTATGACCACTGCGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCCCTTCGAA	303		
QY	243	CATCAGACCTACTGCGCAGCGCAGCTCCGGGAGATCCAGATCTCTGCTGGCTTCGCGCAT	302		
DB	304	CATCAGACCTACTGCGCAGCGCAGCTCCGGGAGATCCAGATCTCTGCTGGCTTCGCGCAT	363		
QY	303	GAGAATGTCATCGGCATCCGAGACATTTCTGCGGGCGTCCACCTTGGAAAGCCATGAGAGAT	362		
DB	364	GAGAATGTCATCGGCATCCGAGACATTTCTGCGGGCGTCCACCTTGGAAAGCCATGAGAGAT	423		
QY	363	GTCTACATTTGTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG	422		
DB	424	GTCTACATTTGTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG	483		
QY	423	CTGAGCAATGACCATATCTGTACTTTCCTCTACAGATCTCTGCGGGGCTCAAGTACATC	482		
DB	484	CTGAGCAATGACCATATCTGTACTTTCCTCTACAGATCTCTGCGGGGCTCAAGTACATC	543		
QY	483	CACTCCGCCAAGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGTCTCATCAACACCACC	542		
DB	544	CACTCCGCCAAGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGTCTCAGCAACACCACC	603		
QY	543	TGCGACCTTAAGATTTGTGATTTGCGCTGCGCCCGGATTTGCCGATCTCTGAGCATGACAC	602		
DB	604	TGCGACCTTAAGATTTGTGATTTGCGCTGCGCCCGGATTTGCCGATCTCTGAGCATGACAC	663		
QY	603	ACCGGCTTCTGACCGGAGTATGTGGCTACGCGTGTGTACCGGGCCCGAGAGATCATGCTG	662		
DB	664	ACCGGCTTCTGACCGGAGTATGTGGCTACGCGTGTGTACCGGGCCCGAGAGATCATGCTG	723		
QY	663	AACTCCAAGGGCTATACCAAGTCCATCGACATCTGTCTGTGGGTGCAATTTCTGGGCTGAG	722		
DB	724	AACTCCAAGGGCTATACCAAGTCCATCGACATCTGTCTGTGGGTGCAATTTCTGGGCTGAG	783		
QY	723	ATGCTCTCTAAACCGGGCCCATCTTCCCTGGCAAGCACTACTCGGATCAAGTCAACCAATT	782		
DB	784	ATGCTCTCTAAACCGGGCCCATCTTCCCTGGCAAGCACTACTCGGATCAAGTCAACCAATT	843		
QY	783	CTGGGATCTCTGGGCTCCCATCCAGGAGGACTGAAATTTGATATCATCAACATGAGGCC	842		
DB	844	CTGGGATCTCTGGGCTCCCATCCAGGAGGACTGAAATTTGATATCATCAACATGAGGCC	903		
QY	843	CGAAACTACTACAGTCTCTGCCCTCCAAAGTGGCTTGGGCCAAGCTTTTCCCC	902		
DB	904	CGAAACTACTACAGTCTCTGCCCTCCAAAGTGGCTTGGGCCAAGCTTTTCCCC	963		
QY	903	AAGTCAGACTCCAAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAAACCCCAATAA	962		

964	Db	 AAGTCAAGCTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAACCCCAATAAA	1023
963	Qy	CGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACTCTGGAGCAGTACTATGACCCGACG	1022
1024	Db	CGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACTCTGGAGCAGTACTATGACCCGACG	1083
1023	Qy	GATGAGCCAGTGGCCGAGGAGCCCTTCACTTTGGCCATGGAGCTGGATGACCTTACCTTAAG	1082
1084	Db	GATGAGCCAGTGGCCGAGGAGCCCTTCACTTTGGCCATGGAGCTGGATGACCTTACCTTAAG	1143
1083	Qy	GAGGGCTGAAGAGTCACTTTCCAGGAGACAGCAGCTTCCAGGCCGCGAGTGCTGGAG	1142
1144	Db	GAGGGCTGAAGAGTCACTTTCCAGGAGACAGCAGCTTCCAGGCCGCGAGTGCTGGAG	1203
1143	Qy	GCCTTCCTAGGCCACAGACAGATCTCTGCACCCCTGGGGCTTGGAAACGAG	1202
1204	Db	GCCTTCCTAGGCCACAGACAGATCTCTGCACCCCTGGGGCTTGGAAACGAG	1243
1203	Qy	GCAAGAGGTCACTGAGGGCTCTGTCAACCAAGACCTGCTCTCTGCTGCCCTCTCCCG	1262
1244	Db	-----GGACCTGCTCTCTGCTGCCCTCTCCCG	1272
1263	Qy	CCAGACTGTTAGAAAATGGACACTGTGCCAGGCCGAGACCTTGGCAGCCCAAGCCCGGGT	1322
1273	Db	CCAGACTGTTAGAAAATGGACACTGTGCCAGGCCGAGACCTTGGCAGCCCAAGCCCGGGT	1332
1323	Qy	GGAGCATGGGCTGGGCCACCTCTCTCTTTGCTGAGGCTTCAAGCTTCAAGGAGGCCAAG	1382
1333	Db	GGAGCATGGGCTGGGCCACCTCTCTCTTTGCTGAGGCTTCAAGCTTCAAGGAGGCCAAG	1392
1383	Qy	GCCTTCTCTCCCAACCGGCTCCCAACCGGGGCTCGGAGCTCAAGTGGCCCAAGTTC	1442
1393	Db	GCCTTCTCTCCCAACCGGCTCCCAACCGGGGCTCGGAGCTCAAGTGGCCCAAGTTC	1452
1443	Qy	AATCTCCC---GCTGCTGCTGCTGCGCCCTTACCTTCCAGAGGCTCCAGTCTCTGGCAG	1499
1453	Db	AATCTCCGCTGCTGCTGCTGCTGGGCCCTTACCTTCCCAAGGCTCCAGTCTCTGGCAG	1512
1500	Qy	TTCTGGAATGGAAGGGTTCTGGCTGCCCCCAACCTGTCTGAAGGGCAGAGGTGGAGTGGG	1559
1513	Db	TTCTGGAATGGAAGGGTTCTGGCTGCCCCCAACCTGTCTGAAGGGCAGAGGTGGAGTGGG	1572
1560	Qy	GGGGCTCAGTGAAGGACTCAAGGGCCATGCTGCCCCCTCATCTCATTTCAAAACCCACCC	1619
1573	Db	GGGGCTCAGTGAAGGACTCAAGGGCCATGCTGCCCCCTCATCTCATTTCAAAACCCACCC	1632
1620	Qy	TAGTTTCCCTGAAGGAAATTTCTTAGTCTCAAGGGCTAGCATCTCTGAGGAGCAGGCC	1679
1633	Db	TAGTTTCCCTGAAGGAAATTTCTTAGTCTCAAGGGCTAGCATCTCTGAGGAGCAGGCC	1692
1680	Qy	GGGGCGAATCCCTTCCCTGTCAAAAGCTGTCACTTGGCGTGGCTTCTGTGTGTGT	1739
1693	Db	GGGGCGAATCCCTTCCCTGTCAAAAGCTGTCACTTGGCGTGGCTTCTGTGTGTGT	1752
1740	Qy	GCTCAGCAGAAGTGGAGCTGGGGGGGCTGGAGAGCCGGGCCCTGCCACCTCCCTGCAC	1799
1753	Db	GCTCAGCAGAAGTGGAGCTGGGGGGGCTGGAGAGCCGGGCCCTGCCACCTCCCTGCAC	1812
1800	Qy	CCGTCTAATATATAAATATAGAGATGTCTATGGCTG	1837
1813	Db	CCGTCTAATAATAAATATAGAGATGTCTATGGCTG	1850

RESULT 4			
LOCUS	CS063502	1866 bp	DNA
DEFINITION	Sequence 486 from Patent EP1522594.		linear
ACCESSION	CS063502		
VERSION	CS063502.1		
KEYWORDS	. GI:62816804		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

[illegible]

Db	844	CTGGGCATCTCTGGGCTCCCCATCCAGGAGGACCTGAAATTGTATCATCAACATGAAGGCC	903
Qy	843	CGAAACTACCTACAGTCTCTGCCCCTCAAGACCAAGGTGCTTGGGCAAGCTTTTCCCCC	902
Db	904	CGAAACTACCTACAGTCTCTGCCCCTCAAGACCAAGGTGCTTGGGCAAGCTTTTCCCCC	963
Qy	903	AAGTCAGACTCCAAAGCCCTTGACCTCTGGACCGGATGTTAACTTTAAACCCCAATAAA	962
Db	964	AAGTCAGACTCCAAAGCCCTTGACCTCTGGACCGGATGTTAACTTTAAACCCCAATAAA	1023
Qy	963	CGGATCAAGTGAGAGGAGCGCTGGCTCACCCCTTACTGTGAGCAGTACTATGACCGGAGC	1022
Db	1024	CGGATCAAGTGAGAGGAGCGCTGGCTCACCCCTTACTGTGAGCAGTACTATGACCGGAGC	1083
Qy	1023	GATGAGCAGTGGCCGAGGAGCCCTTCACTTTGCGCATGAGCTGGATGACCTTACCTAAG	1082
Db	1084	GATGAGCAGTGGCCGAGGAGCCCTTCACTTTGCGCATGAGCTGGATGACCTTACCTAAG	1143
Qy	1083	GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCGTTCCAGCCCGGAGTGTCTGGAG	1142
Db	1144	GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCGTTCCAGCCCGGAGTGTCTGGAG	1203
Qy	1143	GCCCCCTAGCCACAGACAGATCTCTGCAACCTTGGGCGCTTGGAACAGAACTGGCAAGAG	1202
Db	1204	GCCCCCTAGCCACAGACAGATCTCTGCAACCTTGGGCGCTTGGAACAGAACTGGCAAGAG	1243
Qy	1203	GCAAGAGTCACTGAGGGCCTCTGTCAACCAGAGACCTGCTCTCTGCTGCGCCCTCTCCCG	1262
Db	1244	-----GGACCTGCTCTCTGCTGCGCCCTCTCCCG	1272
Qy	1263	CCAGACTGTTAGAAAAATGGACACTGTGCCAGCCCGGACCTTGGCAGCCAGAGCCCGGGT	1322
Db	1273	CCAGACTGTTAGAAAAATGGACACTGTGCCAGCCCGGACCTTGGCAGCCAGAGCCCGGGT	1332
Qy	1323	GGAGCATGGGCGCTGGCCACCTCTCTCTCTTTGCTGAGGCGCTCCAGCTTTCAGGAGGCAAG	1382
Db	1333	GGAGCATGGGCGCTGGCCACCTCTCTCTCTTTGCTGAGGCGCTCCAGCTTTCAGGAGGCAAG	1392
Qy	1383	GCCTTCTCTCCACACCGCGCCCTCCACAGGGGCGCTCGGAGCTCAGGTGGCCCCAGTTC	1442
Db	1393	GCCTTCTCTCCACACCGCGCCCTCCACAGGGGCGCTCGGAGCTCAGGTGGCCCCAGTTC	1452
Qy	1443	AATCTCTCC--GCCTGCTGCTGCGGCCCTTACCTTCCACAGGCTCCAGTCTCTGGCAG	1499
Db	1453	AATCTCTCGCTGCTGCTGCTGCGGCCCTTACCTTCCACAGGCTCCAGTCTCTGGCAG	1512
Qy	1500	TTCTGGAATGGAAGGGTTCTGGCTGCCCCCAACCTGTGAAAGGCGAGGTGGAGGGTGGG	1559
Db	1513	TTCTGGAATGGAAGGGTTCTGGCTGCCCCCAACCTGTGAAAGGCGAGGTGGAGGGTGGG	1572
Qy	1560	GGGCGCTGAGTAGGACTCAGGGGCATGCTGCCCCCTCATCTCATTTCAAACCCCAACC	1619
Db	1573	GGGCGCTGAGTAGGACTCAGGGGCATGCTGCCCCCTCATCTCATTTCAAACCCCAACC	1632
Qy	1620	TAGTTTCCCTGGAAGGAACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGAGCCAGGCC	1679
Db	1633	TAGTTTCCCTGGAAGGAACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGAGCCAGGCC	1692
Qy	1680	GGGCGGAATCCCTCTCCCTGTCAAAAGCTGTCACTTTCGGGTGCGCTCGCTGTCTGTGTGT	1739
Db	1693	GGGCGGAATCCCTCTCCCTGTCAAAAGCTGTCACTTTCGGGTGCGCTCGCTGTCTGTGTGT	1752
Qy	1740	GGTGAGCAGAGTGGAGCTGGGGGGCGTGGAGAGCCCGGCGCCCTGCGACCTCCCTGAC	1799
Db	1753	GGTGAGCAGAGTGGAGCTGGGGGGCGTGGAGAGCCCGGCGCCCTGCGACCTCCCTGAC	1812
Qy	1800	CCGCTCAATATATAAATATAGAGATGTCTATGGCTG	1837
Db	1813	CCGCTCAATATATAAATATAGAGATGTCTATGGCTG	1850
RESULT 5			
CS081283			


```
QY 723 ATGCTCTTAACGGCCCATCTTCTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 782
Db 784 ATGCTCTTAACGGCCCATCTTCTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 843
QY 783 CTGGGATCTCTGGCTCCCATCCAGGAGGAGCTGAATTGTATCATCAACATGAAGGCC 842
Db 844 CTGGGATCTCTGGCTCCCATCCAGGAGGAGCTGAATTGTATCATCAACATGAAGGCC 903
QY 843 CGAAACTACCTACAGTCTCTGCGCTCCCAAGCAAGTGGCTGGGCCAAGCTTTTCCCC 902
Db 904 CGAAACTACCTACAGTCTCTGCGCTCCCAAGCAAGTGGCTGGGCCAAGCTTTTCCCC 963
QY 903 AAGTCAGACTCCAAAGCCCTTGACTCTGCTGGACCGGATGTAACCTTTAAACCCCAATAA 962
Db 964 AAGTCAGACTCCAAAGCCCTTGACTCTGCTGGACCGGATGTAACCTTTAAACCCCAATAA 1023
QY 963 CGATACAGTGAAGGAGCGCTGGCTCACCCTTACTGGAGCAGTACTATGACCCGAGC 1022
Db 1024 CGATACAGTGAAGGAGCGCTGGCTCACCCTTACTGGAGCAGTACTATGACCCGAGC 1083
QY 1023 GATGAGCAGTGGCCGAGGAGCCCTTCCACTTCGCGATGAGCTGGATGACCTACCTAAG 1082
Db 1084 GATGAGCAGTGGCCGAGGAGCCCTTCCACTTCGCGATGAGCTGGATGACCTACCTAAG 1143
QY 1083 GAGCGGCTGAAGGAGTCTATCTTCCAGGAGACGACGCTTCCAGCCCGGAGTGTGGAG 1142
Db 1144 GAGCGGCTGAAGGAGTCTATCTTCCAGGAGACGACGCTTCCAGCCCGGAGTGTGGAG 1203
QY 1143 GCCCCTTAGCCCAAGACAGATCTCTGACACCTGCGGCGCTTGGAAACAGAAAGAG 1202
Db 1204 GCCCCTTAGCCCAAGACAGATCTCTGACACCTGCGGCGCTTGGAAACAGAAAGAG 1243
QY 1203 GCAAGAGTCACTGAGGGCCTCTGTACCCAGAGACCTGCGCTCTGCTGCGCCCTCTCCCG 1262
Db 1244 -----GACCTGCTCTCTGCTGCGCCCTCTCTCCCG 1272
QY 1263 CCAGACTGTTAGAAAATGGACACTGTGCCCAGCCCGGACCTTGGCAGCCAGCCCGGGGT 1322
Db 1273 CCAGACTGTTAGAAAATGGACACTGTGCCCAGCCCGGACCTTGGCAGCCAGCCCGGGGT 1332
QY 1323 GAGAGATGGGCTGGCCACCTCTCTCTCTTGTGAGGCTTCCAGCTTCAGGAGGCGCAG 1382
Db 1333 GAGAGATGGGCTGGCCACCTCTCTCTCTTGTGAGGCTTCCAGCTTCAGGAGGCGCAG 1392
QY 1383 GCCTTCTCTCCACCCGCGCTCCCAAGGCGCTCGGAGCTCAGGTGGCCCGAGTTC 1442
Db 1393 GCCTTCTCTCCACCCGCGCTCCCAAGGCGCTCGGAGCTCAGGTGGCCCGAGTTC 1452
QY 1443 AATCTCCC---GCTGCTGCTGCTGCGCCCTTACCTTCCCGAGGCTCCAGTCTCTGGCAG 1499
Db 1453 AATCTCCGCTGCTGCTGCTGCGCCCTTACCTTCCCGAGGCTCCAGTCTCTGGCAG 1512
QY 1500 TTCTGGAATGGAAGGTTCTGGCTGCCCCCAACTCTGTGAAGGGGAGAGTGGAGGGTGGG 1559
Db 1513 TTCTGGAATGGAAGGTTCTGGCTGCCCCCAACTCTGTGAAGGGGAGAGTGGAGGGTGGG 1572
QY 1560 GGGCGCTAGTAGGACTCAGGGGCGATGCTGCGCCCTCATCTCATTTCAAAACCCACCC 1619
Db 1573 GGGCGCTAGTAGGACTCAGGGGCGATGCTGCGCCCTCATCTCATTTCAAAACCCACCC 1632
QY 1620 TAGTTTCCCTGAAGGAATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1679
Db 1633 TAGTTTCCCTGAAGGAATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1692
QY 1680 GGGCCGAATCCCTCTCTGTAAGAGTGTCACTTTCGGGTGCGCTCGTCTCTGTTGTGT 1739
Db 1693 GGGCCGAATCCCTCTCTGTAAGAGTGTCACTTTCGGGTGCGCTCGTCTCTGTTGTGT 1752
QY 1740 GGTGAGCAGAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTCGCCACCTCCCTGAC 1799
Db 1753 GGTGAGCAGAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTCGCCACCTCCCTGAC 1812
```

```
QY 1800 CGTCTAATATATAATATAGAGATGTGTCTATGGCTG 1837
Db 1813 CGGTCTAATATATAATATAGAGATGTGTCTATGGCTG 1850
RESULT 6
LOCUS AX770530 1866 bp DNA linear PAT 02-JUL-2003
DEFINITION Sequence 36 from Patent WO03031650.
ACCESSION AX770530
VERSION AX770530.1 GI:32437906
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Munnes, M., Gehrman, M., Wick, M. and Schmitz, G.
TITLE Genes and proteins for prevention, prediction, prognosis and
therapy of cardiovascular disease
JOURNAL Patent: WO 03031650-A 36 17-APR-2003;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..1866
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 93.0%; Score 1708.4; DB 6; Length 1866;
Best Local Similarity 97.0%; Pred. No. 7e-290; 1; Indels 54; Gaps 2;
Matches 1783; Conservative 0; Mismatches 0;
QY 3 GGAGTGGAGATGGCGCGCGGCTCAGGGGGGGGGGGGGGGGAGAGCCCGTAGAAC 62
Db 64 GGAGTGGAGATGGCGCGCGGCGGCTCAGGGGGGGGGGGGGGGGAGAGCCCGTAGAAC 123
QY 63 GAGGGGGTGGCGCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTGAC 122
Db 124 GAGGGGGTGGCGCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCGGTTGAC 183
QY 123 GTGGGGCGCGGCTTACACGAGTTCAGTACATCGGGAGGGCGGTACGGCATGGTCAGC 182
Db 184 GTGGGGCGCGGCTTACACGAGTTCAGTACATCGGGAGGGCGGTACGGCATGGTCAGC 243
QY 183 TCGGCTTATGACCACTGGCGCAAGACTCGCTGGCCATCAAGAAGATCAGCCCTTCGAA 242
Db 244 TCGGCTTATGACCACTGGCGCAAGACTCGCTGGCCATCAAGAAGATCAGCCCTTCGAA 303
QY 243 CATCAGACCTTACTGCCAGCGCAGCTCCGGGAGATCCAGATCTGTGGCTTCGGCAT 302
Db 304 CATCAGACCTTACTGCCAGCGCAGCTCCGGGAGATCCAGATCTGTGGCTTCGGCAT 363
QY 303 GAGATGTCTATCGGCATTCGAGACATTCGCGGGCGTCCACCTCGAAGCCATGAGAGAT 362
Db 364 GAGATGTCTATCGGCATTCGAGACATTCGCGGGCGTCCACCTCGAAGCCATGAGAGAT 423
QY 363 GTCTACATTTGTGAGAGCCTGTATGGAGACTGACTGTACAAAGTTGCTGAAAGCCAGCAG 422
Db 424 GTCTACATTTGTGAGAGCCTGTATGGAGACTGACTGTACAAAGTTGCTGAAAGCCAGCAG 483
QY 423 CTGAGCAATGACCATATCTGCTACTTCTTACAGATCTCGGGGGCTCAAGTACATC 482
Db 484 CTGAGCAATGACCATATCTGCTACTTCTTACAGATCTCGGGGGCTCAAGTACATC 543
QY 483 CATCTCGCCAACTGTCTCCACCGAGATCTAAGCCCTCCAGCTGCTCATCACACCC 542
Db 544 CATCTCGCCAACTGTCTCCACCGAGATCTAAGCCCTCCAGCTGCTCATCACACCC 603
QY 543 TGGAGCTTAAAGATTTGTGATTTTCGGCTCGCCCGGATTCGCCATCTGTAGAGCATGACCAC 602
Db 604 TGGAGCTTAAAGATTTGTGATTTTCGGCTCGCCCGGATTCGCCATCTGTAGAGCATGACCAC 663
```

603 ACCGGCTTCTGACGAGTATGTGGCTACGCGTGTGTACCGGGCCCCAGAGATCATGCTG 662
664 ACCGGCTTCTGACGAGTATGTGGCTACGCGTGTGTACCGGGCCCCAGAGATCATGCTG 723
663 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTTCTGGCTGAG 722
724 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTTCTGGCTGAG 783
723 ATGCTCTTAAACGGGCCCATCTTCTCGTGAAGCACTACCTGGATCAGCTCAACCAATTT 782
784 ATGCTCTTAAACGGGCCCATCTTCTCGTGAAGCACTACCTGGATCAGCTCAACCAATTT 843
783 CTGGGATCTCTGGGCTCCCATCCAGGAGGACCTGAATTTGATCATCAACATGAGGCC 842
844 CTGGGATCTCTGGGCTCCCATCCAGGAGGACCTGAATTTGATCATCAACATGAGGCC 903
843 CGAAACTTACTACAGTCTCTGCGCTCCAAAGCAACAGGTGGCTTGGGCCAAGCTTTTCCCC 902
904 CGAAACTTACTACAGTCTCTGCGCTCCAAAGCAACAGGTGGCTTGGGCCAAGCTTTTCCCC 963
903 AAGTCAGATCTCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAACCCCAATAAA 962
964 AAGTCAGATCTCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAACCCCAATAAA 1023
963 CGGATCACAGTGGAGAGCGCTGGCTCACCCCTACCTGAGCAGTACTATGACCCGAGC 1022
1024 CGGATCACAGTGGAGAGCGCTGGCTCACCCCTACCTGAGCAGTACTATGACCCGAGC 1083
1023 GATGACCATGTCGCGAGGAGCCCTTTCAGCTTGGCCATGAGCTGATGACCTACCTAAG 1082
1084 GATGACCATGTCGCGAGGAGCCCTTTCAGCTTGGCCATGAGCTGATGACCTACCTAAG 1143
1083 GAGCGCTGAAGAGTCTATTTCCAGGAGACAGACGCTTTCAGCCCGAGGTGCTGGAG 1142
1144 GAGCGCTGAAGAGTCTATTTCCAGGAGACAGACGCTTTCAGCCCGAGGTGCTGGAG 1203
1143 GCGCCCTAGCCAGACAGACATCTCTGACCTTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1202
1204 GCGCCCTAGCCAGACAGACATCTCTGACCTTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1243
1203 GCAAGAGGTCACTGAGGGCTCTGTGTCACGAGGACCTGCTGCTGCTGCTGCTGCTGCTG 1262
1244 -----GGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1272
1263 CCAGACTGTTAGAAATGACACTGTGCGCAGCCCGGACCTTGGCAGCCAGCGCGGGGT 1322
1273 CCAGACTGTTAGAAATGACACTGTGCGCAGCCCGGACCTTGGCAGCCAGCGCGGGGT 1332
1323 GGAGCATGGGCTGGCAGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
1333 GGAGCATGGGCTGGCAGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
1383 GCCTTCTCTCCACCGCCCTCCCAAGGGGCTCGGAGCTCAGGTGGGCCCGAGTTTC 1442
1393 GCCTTCTCTCCACCGCCCTCCCAAGGGGCTCGGAGCTCAGGTGGGCCCGAGTTTC 1452
1443 AATCTCCC---GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1499
1453 AATCTCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1512
1500 TTCTGAATGGAGGGTCTTGGGCTGCGCCCAACTGCTGAGGGCAGAGGTGGGGTGGG 1559
1513 TTCTGAATGGAGGGTCTTGGGCTGCGCCCAACTGCTGAGGGCAGAGGTGGGGTGGG 1572
1560 GGGCGTGTAGTGGGACTCAGGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
1573 GGGCGTGTAGTGGGACTCAGGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
1620 TAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1679
1633 TAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1692

1680 GGGCCGAATCCCTCCCTGTCACAAAGCTGTCACTTCCGCTGCGCTCGCTGCTCTCTGTTGT 1739
1693 GGGCCGAATCCCTCCCTGTCACAAAGCTGTCACTTCCGCTGCGCTCGCTGCTCTGTTGT 1752
1740 GGTGAGCAGAAAGTGTGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTGCACCTCCCTGAC 1799
1753 GGTGAGCAGAAAGTGTGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTGCACCTCCCTGAC 1812
1800 CGTCTAATATATAATATAGAGATGCTCTATGGCTG 1837
1813 CGCTAATATATAATATAGAGATGCTCTATGGCTG 1850

RESULT 7
HSERK1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Human ERK1 mRNA for protein serine/threonine kinase.
X60188.1 GI:31220
erk1 gene; protein-serine/threonine kinase.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 1866)
Charest, D.L., Mordret, G., Harder, K.W., Jirik, F. and Pelech, S.L.
Molecular cloning, expression, and characterization of the human
mitogen-activated protein kinase p46erk1
Mol. Cell. Biol. 13 (8), 4679-4690 (1993)
7687743

2 (bases 1 to 1866)
Pelech, S.L.
Direct Submission
Submitted (23-JUN-1991) S.L. Pelech, Biomedical Res Centre, 2222
Health Science Hall, Univ of British Columbia, Vancouver B C V6T
1Z3, CANADA

FEATURES
source
1..1866
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="p26a-Beta-3"
/cell_line="HepG2"
/cell_type="hepatoblastoma"
/tissue_type="liver tumor"
/clone_lib="HepG2"
1..1866
/gene="ERK1"
1..1866
/gene="ERK1"
73..1212
/evidence=experimental
/gene="ERK1"
/codon_start=1
/product="protein serine/threonine kinase"
/protein_id="CAA42744.1"
/db_xref="GI:31221"
/db_xref="GOA:P27361"
/db_xref="UniProt/Swiss-Prot:P27361"
/translation="MAAAGAGGGGPRTEGVGPVGEVEMVKGPFDVGPRTYQ
LQYIGEGAVGMVSSAYDHVRTRVAIKKISPEHOTYCORLTREIQILRRFHNVI
IRDLIRASTLEMRDVIYQDLMDLYKLLKSOQLSNDHLCYFLYQILRGKYIHS
NVLKDLKPSNLLSNITCDLKDFGLARIADPEHDHTGFLTEYVATRYRAPEIMLN
SKGYTKSDIWSVGCILAEMLSNRPFPGKHYLDQNLHILGLISGFSQEDLNCIMK
ARNYLQSLPSKTKVAWAKLFPKSDSKALDLDRLMTFNPKNRITVTEALHPYLSQY
DPTDEPVAEEPTFAMELDLPLKERLKELIFQETARFQGVLEAP"

gene
mRNA
CDS

Query Match 93.0%; Score 1708.4; DB 8; Length 1866;
Best Local Similarity 97.0%; Pred. No. 7e-290;
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;

CDS		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
		73..1212	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD83831.1"	
		/db_xref="GI:29334001"	
		/translation="MAAATAAQGGGGEPRTGVPVGEVEMVKGPDPVGPRTQ	
		LQYIGGATVGMSSAYDHWKRVKAIKISPFHQYICORTUREIQILLRPHENVIG	
		IRDIIRASTLEARDYIVQDLMETDLYLKLSQQLSNDHICYFIYQILRGDKYIHS	
		NVLRLKPLNSLNTCTDLKI CDFGLRIADPEHDHITGLFTEYVATWYRAPEIMLN	
		SVGYTKSIDWSVGTILAMLSNRPIFPKGHVLDOLNHLILGILGSPQSDLNICINMK	
		ARNYLQSLPSKTKVAWAKLPKSDSKALDLLRMLTFENPKRITVEEALAHPLYLQYY	
		DTDFEVAPEPTFMELDDLPKRLKELIFQETARFQPVLEAP"	
ORIGIN		Query Match 93.0%; Score 1708.4; DB 6; Length 3732;	
		Best Local Similarity 97.0%; Pred. No. 6.5e-290;	
		Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;	
QY	3	GGAGTGGAGATGGCGCGCGGCTCAGGGGGCGGGGGGAGCCCCGTAGAACC 62	
DB	64	GGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGGGAGCCCCGTAGAACC 123	
QY	63	GAGGGGTGGCCCCGGGGTCCCGGGGAGGTGGAGTGGTGAAGGGGAGCGGTTCCGAC 122	
DB	124	GAGGGGGTGGCCCCGGGGTCCCGGGGAGGTGGAGTGGTGAAGGGGAGCGGTTCCGAC 183	
QY	123	GTGGGCGCGCGCTACACGACATGTCAGTATCATCGGCGAGGGCGGTACGGCATGGTCAGC 182	
DB	184	GTGGGCGCGCGCTACACGACATGTCAGTATCATCGGCGAGGGCGGTACGGCATGGTCAGC 243	
QY	183	TCGGCCTATGACCAACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATACGCCCTTCGAA 242	
DB	244	TCGGCCTATGACCAACGTGCGCAAGACTCGCGTGGCCATCAAGAAGATACGCCCTTCGAA 303	
QY	243	CATCAGACCTACTGCGCAGCGACGCTCCGGGAGATCCAGATCTGCTGCGCTTCGGCCAT 302	
DB	304	CATCAGACCTACTGCGCAGCGACGCTCCGGGAGATCCAGATCTGCTGCGCTTCGGCCAT 363	
QY	303	GAGAAATGTCATCGGCATCCGAGACATTCGCGGGCTCCACCTCGGAAGCCATGAGAGAT 362	
DB	364	GAGAAATGTCATCGGCATCCGAGACATTCGCGGGCTCCACCTCGGAAGCCATGAGAGAT 423	
QY	363	GTCTACATTTGTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG 422	
DB	424	GTCTACATTTGTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG 483	
QY	423	CTGAGCAATGACCATATCTGCTACTTCTCTACCATCATCTGCGGGGCTCAAGTACATC 482	
DB	484	CTGAGCAATGACCATATCTGCTACTTCTCTACCATCATCTGCGGGGCTCAAGTACATC 543	
QY	483	CACCTCCGCCAAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACACC 542	
DB	544	CACCTCCGCCAAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCAGCAACACCAC 603	
QY	543	TGCGACCTTAAGATTTGTGATTTGCGCTGCGCCCGGATTCGCCATCTCTGAGCATGACAC 602	
DB	604	TGCGACCTTAAGATTTGTGATTTGCGCTGCGCCCGGATTCGCCATCTCTGAGCATGACAC 663	
QY	603	ACCGGCTTCTGACCGAGATATGGCTACGCGTGGTACCGGGCCCGAGAGATCATGCTG 662	
DB	664	ACCGGCTTCTGACCGAGATATGGCTACGCGTGGTACCGGGCCCGAGAGATCATGCTG 723	
QY	663	AACTCCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCAATCTTGGCTGAG 722	
DB	724	AACTCCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCAATCTTGGCTGAG 783	
QY	723	ATGCTCTCTAAACCGGGCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 782	
DB	784	ATGCTCTCTAAACCGGGCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATT 843	

QY	783	CTGGGCATCTTGGGCTCCCATCCAGGAGGACCTGAAATGTATCATCAACATGAAGGCC 842	
DB	844	CTGGGCATCTTGGGCTCCCATCCAGGAGGACCTGAAATGTATCATCAACATGAAGGCC 903	
QY	843	CGAAACTACCTACGTCTCTGCCCCTCAAGACCAAGGTGGCTTGGGCGAGCTTTTCCCC 902	
DB	904	CGAAACTACCTACGTCTCTGCCCCTCAAGACCAAGGTGGCTTGGGCGAGCTTTTCCCC 963	
QY	903	AGTTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAA 962	
DB	964	AGTTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAA 1023	
QY	963	CGGATCACAAGTGGAGGAAGCGCTGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG 1022	
DB	1024	CGGATCACAAGTGGAGGAAGCGCTGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG 1083	
QY	1023	GATGAGCAGTGGCGGAGAGCCCTTACCTTTCGCTGAGAGCTGGAGTGGATCCTACTAAG 1082	
DB	1084	GATGAGCAGTGGCGGAGAGCCCTTACCTTTCGCTGAGAGCTGGAGTGGATCCTACTAAG 1143	
QY	1083	GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1142	
DB	1144	GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1203	
QY	1143	GCCCCCTAGCCACAGACACATCTCTGCAACCCTGGGGGCTCGAAACAGAACTGGCAAGAG 1202	
DB	1204	GCCCCCTAGCCACAGACACATCTCTGCAACCCTGGGGGCT----- 1243	
QY	1203	GCAAGAGTCACTGAGGGGCTCTGTACCCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1262	
DB	1244	-----GGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1272	
QY	1263	CCAGACTGTTAGAAATGGACACTGTGCCCAGCGACCTTTGGCAGCCAGCCAGCGGGGT 1322	
DB	1273	CCAGACTGTTAGAAATGGACACTGTGCCCAGCGACCTTTGGCAGCCAGCCAGCGGGGT 1332	
QY	1323	GGAGCATGGGCTGGCCACCTCTCTCTTTGCTGAGGGCTCCAGCTTCAGGCAGGCGCAAG 1382	
DB	1333	GGAGCATGGGCTGGCCACCTCTCTCTTTGCTGAGGGCTCCAGCTTCAGGCAGGCGCAAG 1392	
QY	1383	GCCTTCTCTCCCAACCGGCTTCCCAACGGGGCTTCGGGAGCTCAGTGGGCGCCAGTTTC 1442	
DB	1393	GCCTTCTCTCCCAACCGGCTTCCCAACGGGGCTTCGGGAGCTCAGTGGGCGCCAGTTTC 1452	
QY	1443	AAATCTCCC---GCTGCTGCTGCTGCGCCCTTACCTTCCCAGCGTCCAGTCTCTTGGCAG 1499	
DB	1453	AAATCTCCCCTGCTGCTGCTGCGCCCTTACCTTCCCAGCGTCCAGTCTCTTGGCAG 1512	
QY	1500	TTCTGGAATGGAGGGTTCTGGCTGCCCCCAACCTGCTGAAGGGGAGAGGTGAGGGTGGG 1559	
DB	1513	TTCTGGAATGGAGGGTTCTGGCTGCCCCCAACCTGCTGAAGGGGAGAGGTGAGGGTGGG 1572	
QY	1560	GGGCGCTGATGAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAACCCACCC 1619	
DB	1573	GGGCGCTGATGAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAACCCACCC 1632	
QY	1620	TAGTTTCCCTGAAGGAACATTTCTTAGTCTCAAGGCTAGCATCCTCTGAGGAGCCAGGCC 1679	
DB	1633	TAGTTTCCCTGAAGGAACATTTCTTAGTCTCAAGGCTAGCATCCTCTGAGGAGCCAGGCC 1692	
QY	1680	GGGCGGAATCCCTCTCTGCTCAAAGCTGTCACCTGCGTGGCCCTCGCTGCTTCTGTGTGT 1739	
DB	1693	GGGCGGAATCCCTCTCTGCTCAAAGCTGTCACCTGCGTGGCCCTCGCTGCTTCTGTGTGT 1752	
QY	1740	GGTGAGCAGAAATGAGAGCTGGGGGGGTGGAGAGCCCGGCGCCCTGCGCACCTCCCTGAC 1799	
DB	1753	GGTGAGCAGAAATGAGAGCTGGGGGGGTGGAGAGCCCGGCGCCCTGCGCACCTCCCTGAC 1812	
QY	1800	CGGCTCTAATATATATATAGAGATGTGTCTATGGCTG 1837	
DB	1813	CGGCTCTAATATATATATATAGAGATGTGTCTATGGCTG 1850	

QY	1551	GAGGTGGGGGGCGCTGAGTACGAGTACAGGCCATGCTGCCCCCTCATCTCATTTCAA	1610
Db	1448	GAGGTGGGGGGCGCTGAGTACGAGTACAGGCCATGCTGCCCCCTCATCTCATTTCAA	1507
QY	1611	ACCCACCCCTAGTTTCCCTGAAGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGG	1670
Db	1508	ACCCACCCCTAGTTTCCCTGAAGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGG	1567
QY	1671	AGCCAGGCGGGCGGAATCCCTCCCTGTCATGAAGCTGTCACCTGCGGTGCTCGTGTGT	1730
Db	1568	AGCCAGGCGGGCGGAATCCCTCCCTGTCATGAAGCTGTCACCTGCGGTGCTCGTGTGT	1627
QY	1731	TCTGTGTGCTGTCAGCAGAGTGGAGCTGGGGGGCTGGAGGCCCGGC-GCCCTGGCCA	1789
Db	1628	TCTGTGTGCTGTCAGCAGAGTGGAGCTGGGGGGCTGGAGGCCCGGCCTGGCCCGCCA	1687
QY	1790	CTCCCTGACCCCTCTAATATATAAATATAGAGATGTGCTATGGCTG	1837
Db	1688	CTCCCTGACCCCTCTAATATATAAATATAGAGATGTGCTATGGCTG	1735
RESULT 10			
AK091009			
LOCUS		2567 bp mRNA linear PRI 30-JAN-2004	
DEFINITION		Homo sapiens CDNA FLJ3690 fis, clone BRAWH2002967, highly similar	
		to MITOGEN-ACTIVATED PROTEIN KINASE 3 (EC 2.7.1.-).	
ACCESSION		AK091009	
VERSION		AK091009.1 GI:21749281	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
		Hominidae; Homo.	
REFERENCE		1	
AUTHORS		Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,	
		Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,	
		Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,	
		Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,	
		Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,	
		Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,	
		Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,	
		Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,	
		Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,	
		Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,	
		Hiraka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,	
		Yosida, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,	
		Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,	
		Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,	
		Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,	
		Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,	
		Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,	
		Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,	
		Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,	
		Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,	
		Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,	
		Ozaki, K., Hiroo, M., Ohmori, Y., Kawabata, A., Hiki, J., Kobatake, N.,	
		Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,	
		Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,	
		Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,	
		Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,	
		Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,	
		Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,	
		Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.	
		Complete sequencing and characterization of 21,243 full-length	
		human cDNAs	
TITLE		Nat. Genet. 36 (1), 40-45 (2004)	
JOURNAL			
PUBMED		14702039	
REFERENCE		2	
AUTHORS		Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,	
		Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,	
		Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,	

		Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,	
		Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,	
		Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K.,	
		Masuno, Y., Nagai, K. and Isogai, T.	
		NEDO human cDNA sequencing project	
TITLE		Unpublished	
REFERENCE		3 (bases 1 to 2567)	
AUTHORS		Isogai, T. and Yamamoto, J.	
JOURNAL		Direct Submission	
		Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7	
		Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan	
		(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)	
COMMENT		NEDO human cDNA sequencing project supported by Ministry of	
		Economy, Trade and Industry of Japan; cDNA full insert sequencing:	
		Research Association for Biotechnology (RAB); cDNA library	
		construction: Helix Research Institute (HRI) (supported by Japan	
		Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,	
		HRI, and Biotechnology Center, National Institute of Technology and	
		Evaluation; clone selection for full insert sequencing: HRI and	
		RAB; annotation: HRI and RAB.	
FEATURES		Location/Qualifiers	
source		1..2567	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="BRAWH2002967"	
		/tissue_type="brain"	
		/clone_lib="BRAWH2"	
		/note="Cloning vector: pME18SFL3"	
ORIGIN			
		Query Match 84.2%; Score 1546.8; DB 8; Length 2567;	
		Best Local Similarity 96.5%; Pred. No. 1.6e-261;	
		Matches 1612; Conservative 0; Mismatches 7; Indels 51; Gaps 1;	
QY	168	TACGGCATGTGTCAGCTCGCGCTATGACCACTGTCGCGTGCAGATCTCGGCGCTCCACCTCG	227
Db	949	TCCGGTCCCTCAGCTCGCGCTATGACCACTGTCGCGTGCAGATCTCGGCGCTCCACCTCG	1008
QY	228	ATCAGCCCTTCGACATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCCCTG	287
Db	1009	ATCAGCCCTTCGACATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCCCTG	1068
QY	288	CTGCGCTTCGCGCATGAGATGTCTATCGGCATCCGAGACATTCGCGGCGCTCCACCTCG	347
Db	1069	CTGCGCTTCGCGCATGAGATGTCTATCGGCATCCGAGACATTCGCGGCGCTCCACCTCG	1128
QY	348	GAAGCCATGAGAGATGTCTATCTGTGCGAGCACTGTGAGAGACTGACTGTACAGTTG	407
Db	1129	GAAGCCATGAGAGATGTCTATCTGTGCGAGCACTGTGAGAGACTGACTGTACAGTTG	1188
QY	408	CTGAAAGCCAGCAGCTGAGCAATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	467
Db	1189	CTGAAAGCCAGCAGCTGAGCAATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1248
QY	468	GGCTCTCAAGTATACCTCCGCTCCGCAACGTCCTCCACCGAGATCTAAAGCCCTCCACCTG	527
Db	1249	GGCTCTCAAGTATACCTCCGCTCCGCAACGTCCTCCACCGAGATCTAAAGCCCTCCACCTG	1308
QY	528	CTCATCAACACCACTCGGACCTTAAAGATTTGTGATTTTGGCTGCGCTGCGCGGATTCGCGAT	587
Db	1309	CTCATCAACACCACTCGGACCTTAAAGATTTTGTGATTTTGGCTGCGCTGCGCGGATTCGCGAT	1368
QY	588	CCTGAGCATGACCAACACCGGCTTCCTGACGGAGTATGTGGCTACCGCTGGTACCGGGCC	647
Db	1369	CCTGAGCATGACCAACACCGGCTTCCTGACGGAGTATGTGGCTACCGCTGGTACCGGGCC	1428
QY	648	CCAGAGATCATCTGAACTCCAAAGGGCTATACCAAGTCCATCGACATCTGTGTGGGC	707
Db	1429	CCAGAGATCATCTGAACTCCAAAGGGCTATACCAAGTCCATCGACATCTGTGTGGGC	1488
QY	708	TCGATTTCTGGCTGAGATGCTCTCTTAACCGCGCCATCTTCCCTGGCAGACATCTACCTGGAT	767

Db	1489	TGCAATTTGGCTGAGATGCTCTCTAAACCGGCCCATCTTCCTCGGCAAGCACTACCTGGAT	1548
Qy	768	CAGCTCAACACATCTTGGGCATCTCTGGGCTCCCTCCATCCAGGAGACCTGAATTGTATC	827
Db	1549	CAGCTCAACACATCTTGGGCATCTCTGGGCTCCCTCCATCCAGGAGACCTGAATTGTATC	1608
Qy	828	ATCAACATGAAGCCCGGAACTACCTACAGTCTCTGCGCTCCAGACCAAGAGTGGCTTGG	887
Db	1609	ATCAACATGAAGCCCGGAACTACCTACAGTCTCTGCGCTCCAGACCAAGAGTGGCTTGG	1668
Qy	888	GCCAGGCTTTTCCCAAGTCAAGTCCAAAGCCCTTACCTGCTGACCGGATGTTAAAC	947
Db	1669	GCCAGGCTTTTCCCAAGTCAAGTCCAAAGCCCTTACCTGCTGACCGGATGTTAAAC	1728
Qy	948	TTTAAACCCCAATAAAGGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAG	1007
Db	1729	TTTAAACCCCAATAAAGGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAG	1788
Qy	1008	TACTATGACCCGACGATAGCAGTGGCGGAGAGCCCTTACCTTGGCCATGGAGCTG	1067
Db	1789	TACTATGACCCGACGATAGCAGTGGCGGAGAGCCCTTACCTTGGCCATGGAGCTG	1848
Qy	1068	GATGACCTACCTAAGGAGCGCTGAAGAGCTCATCTTCCAGGAGCAGCAGCTTCCAG	1127
Db	1849	GATGACCTACCTAAGGAGCGCTGAAGAGCTCATCTTCCAGGAGCAGCAGCTTCCAG	1308
Qy	1128	CCCGAGTGTGGAGGCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCTTGGAC	1187
Db	1909	CCCGAGTGTGGAGGCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCTTGGAC	1963
Qy	1188	AGAACTGGCAAGGCAAGAGGTCACTGAGGGGCTCTGTCAACCCAGGACTGCGCTCTG	1247
Db	1964	-----GGAGCTGCTCCTG	1977
Qy	1248	CTGTGCCCCCTCTCCCGCAGACTGTAGAAAATGACACTGTGTCCAGCCCGGACCTTGGC	1307
Db	1978	CTGTGCCCCCTCTCCCGCAGACTGTAGAAAATGACACTGTGTCCAGCCCGGACCTTGGC	2037
Qy	1308	AGCCAGCGCGGGTGAGAGATGGGCTGGCCACTCTCTCTCTTGTGTAGGCTCCAGC	1367
Db	2038	AGCCAGCGCGGGTGAGAGATGGGCTGGCCACTCTCTCTCTTGTGTAGGCTCCAGC	2097
Qy	1368	TTGAGGAGCCGAGGCTTCTCTCCCAACCCGCTCTCCCAACCCGCGGCTCGGGAGCTC	1427
Db	2098	TTGAGGAGCCGAGGCTTCTCTCCCAACCCGCTCTCCCAACCCGCGGCTCGGGAGCTC	2157
Qy	1428	AGTGGGCCAGATTCATCTCCCGCTGCTGTGTGCGGCTTACCTTCCCGAGCTGCC	1487
Db	2158	AGTGGGCCAGATTCATCTCCCGCTGCTGTGTGCGGCTTACCTTCCCGAGCTGCC	2217
Qy	1488	AGTCTCTGGCAGTTCTGGAATGGAAGGTTCTGGCTGCCCAACCTCTGTAAGGGCAGAG	1547
Db	2218	AGTCTCTGGCAGTTCTGGAATGGAAGGTTCTGGCTGCCCAACCTCTGTAAGGGCAGAG	2277
Qy	1548	GTGAGGAGTGGGGGCGCTGAGTAGGAGACTCAGGGCCATGCTGCCCTCTCATCTCAT	1607
Db	2278	GTGAGGAGTGGGGGCGCTGAGTAGGAGACTCAGGGCCATGCTGCCCTCTCATCTCAT	2337
Qy	1608	CAAAACCCACCTAGTTTCCCTGAAGAACTTCTAGTCTGAAGGCTAGCATCCCTG	1667
Db	2338	CAAAACCCACCTAGTTTCCCTGAAGAACTTCTAGTCTGAAGGCTAGCATCCCTG	2397
Qy	1668	AGGAGCAGCGCGGCGGATCCCTCTCTGTAAGGCTGTCCTTCTGCGTGGCCCTCGCT	1727
Db	2398	AGGAGCAGCGCGGCGGATCCCTCTCTGTAAGGCTGTCCTTCTGCGTGGCCCTCGCT	2457
Qy	1728	GCTTCTGTGTGTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTG	1787
Db	2458	GCTTCTGTGTGTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTG	2517
Qy	1788	CACCTCCCTGACCGCTCTAATATATATATATATATATATATATATATATATATATATAT	1837
Db	2518	CACCTCCCTGACCGCTCTAATATATATATATATATATATATATATATATATATATATAT	2567

RESULT 11
HSM806004

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..2722

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="RZPD:DKFZp68600215"

/db_xref="taxon:9606"

/map="16"

/clone="DKFZp68600215"

/tissue_type="human cervix"

/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host

DH10B; sites SfiIA + SfiIB"

/dev_stage="adult"

<1..2722

/genes="DKFZp68600215"

<960..1931

/genes="DKFZp68600215"

/notes="mitogen-activated protein kinase 3, N-terminus

truncated, not fully spliced"

/codon_start=1

/product="hypothetical protein"

/protein_id="CAD97888.1"

/db_xref="GI:31873913"

/db_xref="GOA:O7Z3H5"

/db_xref="InterPro:IPR000719"

/db_xref="InterPro:IPR002290"

/db_xref="InterPro:IPR003527"

/db_xref="InterPro:IPR008271"

/db_xref="UniProt/TREMBL:O7Z3H5"

/translations="SSAYDHVTRVVAIKKISPFHOTVCORTLREIQILIRPHENV

IGRIDLRASTLEAMRDVTVQDMETDLYKLLKSOQLSNDHCYLYQLRGKYKH

SNVLRDLKPSNLLINTCDLKI CDFGLARIADPHDHTGFLTEYVATWYRAPEIM

LNKSGYTKSIDIVSWGICLAELSNRPFPFGKHYLDQLNHLGILGSPQEDLNCTIN

MKARNYLSQSPSTKVAKWLPKSDSKALDLDRLMTFNPKNRITVVEALHPYLEQ

YDPTDEPVAEBEFTFAMELDDLPKRLKELIFQETARFQGVLEAP"

2559

/gene="DKFZp68600215"

polyA_site

ORIGIN

Query Match

Best Local Similarity

Matches 1606; Conservative

0; Mismatches

7; Indels

51; Gaps

1;

Qy

168 TACGGCATGTGTCAGCTCGGCTATGACACGTCGCGCAAGACTCGCTGGCCATCAAGAG

227

83.9%; Score 1540.8; DB 8; Length 2722;

Pred. No. 1.8e-260;

Matches 1606; Conservative

0; Mismatches

7; Indels

51; Gaps

1;

Db 948 TCCGGTCCCTCAGCTCGGCTATGACACAGTGGCAAGACTCGGTGGCCATCAAGAAG 1007
QY 228 ATCAGCCCTTTCGAACATCAGACCTACTCTGCAGCGCAGCTCGGAGATCCAGATCCTG 287
Db 1008 ATCAGCCCTTTCGAACATCAGACCTACTCTGCAGCGCAGCTCGGAGATCCAGATCCTG 1067
QY 288 CTGCGCTTCGGCATGAGAAATGTCATCGGCATCCGAGACATTCGCGGCGTCCACCCCTG 347
Db 1068 CTGCGCTTCGGCATGAGAAATGTCATCGGCATCCGAGACATTCGCGGCGTCCACCCCTG 1127
QY 348 GAAGCCATCAGAGATGTCATATGTCAGGACCTGATGGAGACTGACCTGTACAAAGTTG 407
Db 1128 GAAGCCATCAGAGATGTCATATGTCAGGACCTGATGGAGACTGACCTGTACAAAGTTG 1187
QY 408 CTGAAAGCCAGCAGCTGAGCAATGACCAATATCTGCTACTTCTTACAGATCCTGGG 467
Db 1188 CTGAAAGCCAGCAGCTGAGCAATGACCAATATCTGCTACTTCTTACAGATCCTGGG 1247
QY 468 GGCCTCAAGTACATCCACTCCGCCACGCTCCACGAGATCTAAAGCCCTCCAAACCTG 527
Db 1248 GGCCTCAAGTACATCCACTCCGCCACGCTCCACGAGATCTAAAGCCCTCCAAACCTG 1307
QY 528 CTATCAACACCACTCGCACTTAAGATTTGTGATTTGCGCTGCGCCCGGATTCGCCAT 587
Db 1308 CTATCAACACCACTCGCACTTAAGATTTGTGATTTGCGCTGCGCCCGGATTCGCCAT 1367
QY 588 CTGAGCATGACACACCGGCTTCCTGACGAGATATGTCGCTACGCGCTGTTACCGGGCC 647
Db 1368 CTGAGCATGACACACCGGCTTCCTGACGAGATATGTCGCTACGCGCTGTTACCGGGCC 1427
QY 648 CCAGAGATCATGCTGAACCTCAAGGCTATACCAAGTCCATCGACATCTGCTGTGGG 707
Db 1428 CCAGAGATCATGCTGAACCTCAAGGCTATACCAAGTCCATCGACATCTGCTGTGGG 1487
QY 708 TGCATTTCTGGCTGAGATGCTCTTAACCGGCCATCTTCCCTGGCAAGCACTACCTGGAT 767
Db 1488 TGCATTTCTGGCTGAGATGCTCTTAACCGGCCATCTTCCCTGGCAAGCACTACCTGGAT 1547
QY 768 CAGCTCAACACATTTCTGGGCTATCTGGGCTCCCATCCAGAGACCTGAATGTATC 827
Db 1548 CAGCTCAACACATTTCTGGGCTATCTGGGCTCCCATCCAGAGACCTGAATGTATC 1607
QY 828 ATCAACATGAAGCCCGGAACTACCTACAGTCTCTGCTCCCTCAAGACCAAGTGGCTGG 887
Db 1608 ATCAACATGAAGCCCGGAACTACCTACAGTCTCTGCTCCCTCCAGACCAAGTGGCTGG 1667
QY 888 GCCAAGCTTTTCCCAAGTCAGACTCCAAAGCCCTTGACCTGCTGCAACCGGATGTTAAC 947
Db 1668 GCCAAGCTTTTCCCAAGTCAGACTCCAAAGCCCTTGACCTGCTGCAACCGGATGTTAAC 1727
QY 948 TTTAACCCCAATAACCGGATCAAGTGGAGGAAAGCGCTGGCTCACCCCTACCTGGAGCAG 1007
Db 1728 TTTAACCCCAATAACCGGATCAAGTGGAGGAAAGCGCTGGCTCACCCCTACCTGGAGCAG 1787
QY 1008 TACTATGACCCGAGATGAGCGAGTGGCGGAGGAGCCCTTACCTTCGCCATGGAGCTG 1067
Db 1788 TACTATGACCCGAGATGAGCGAGTGGCGGAGGAGCCCTTACCTTCGCCATGGAGCTG 1847
QY 1068 GATGACTTACCTAAGGAGCGGCTGAAGGAGCTCATCTTCAGGAGACAGCAGCTTCCAG 1127
Db 1848 GATGACTTACCTAAGGAGCGGCTGAAGGAGCTCATCTTCAGGAGACAGCAGCTTCCAG 1907
QY 1128 CCGGAGTGTGAGGCGCCCTTAGCCACAGACATCTCTGCAACCTGGGCGCTGGAC 1187
Db 1908 CCGGAGTGTGAGGCGCCCTTAGCCACAGACATCTCTGCAACCTGGGCGCTGGAC 1962
QY 1188 AGAAGTGGCAAGAGGCTCACTGAGGCGCTCTGTCAACCCAGGACCTGCTCCCTG 1247
Db 1963 -----GGACCTGCTCCCTG 1976
QY 1248 CTTGCCCTCTCCCGCAGACTGTTAGAAAATGGACATCTGTGCCAGCCGACCTTGGC 1307

Db 1977 CTTGCCCTCTCCCGCAGACTGTTAGAAAATGGACACTGTGTGCCAGGCCGACCTTGGC 2036
QY 1308 AGCCAGGCGCGGGTGGAGCATGGGCTTGCCACCTCTCTCTTTGCTGAGGCTCCACG 1367
Db 2037 AGCCAGGCGCGGGTGGAGCATGGGCTTGCCACCTCTCTCTTTGCTGAGGCTCCACG 2096
QY 1368 TTAGCGAGGCAAGGCTCTCTCTCCCAACCGCCCTCCCAACCGGCGCTCGGAGCTC 1427
Db 2097 TTAGCGAGGCAAGGCTCTCTCTCCCAACCGCCCTCCCAACCGGCGCTCGGAGCTC 2156
QY 1428 AGGTGGCCCACTTCAATCTCCCGCTGCTGTGCTGCGCCCTTACCTTCCCGCAGGCTCC 1487
Db 2157 AGGTGGCCCACTTCAATCTCCCGCTGCTGTGCTGCGCCCTTACCTTCCCGCAGGCTCC 2216
QY 1488 AGTCTCTGCGACTTCTGGAATGGAAGGCTTCTGGCTGCCCAACCTGTGTGAAGGCGCAG 1547
Db 2217 AGTCTCTGCGACTTCTGGAATGGAAGGCTTCTGGCTGCCCAACCTGTGTGAAGGCGCAG 2276
QY 1548 GTGAGGCTGGGGCGCTGAGTAGGGAATCAGGCGCATGCTGCCCGCTCATCTCAT 1607
Db 2277 GTGAGGCTGGGGCGCTGAGTAGGGAATCAGGCGCATGCTGCCCGCTCATCTCAT 2336
QY 1608 CAAACCCACCTAGTTTCCCTGAAAGAACTTCTTAGTCTCAAGGGCTAGCATCCCTG 1667
Db 2337 CAAACCCACCTAGTTTCCCTGAAAGAACTTCTTAGTCTCAAGGGCTAGCATCCCTG 2396
QY 1668 AGGAGCCAGGCGCGGCGGAACTCCCTGCTCAAGCTGTCACTTGGGCTGCCCTCGCT 1727
Db 2397 AGGAGCCAGGCGCGGCGGAACTCCCTGCTCAAGCTGTCACTTGGGCTGCCCTCGCT 2456
QY 1728 GCTTCTGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTG 1787
Db 2457 GCTTCTGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTG 2516
QY 1788 CACCTCCCTGAGCCCTTAATATATATATATATATATATATATATATATATATATAT 1831
Db 2517 CACCTCCCTGAGCCCTTAATATATATATATATATATATATATATATATATATATAT 2560

RESULT 12

HS44KDA
LOCUS H.sapiens 44kDa protein kinase related to rat ERK1. linear PRI 18-APR-2005
DEFINITION Z11696 S38872
ACCESSION Z11696.1 GI:23882
VERSION protein kinase.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 1284)
AUTHORS Gonzalez, F.A., Raden, D.L., Rigby, M.R. and Davis, R.J.
TITLE Heterogeneous expression of four MAP kinase isoforms in human tissues
JOURNAL FEBS Lett. 304 (2-3), 170-178 (1992)

REFERENCE 2 (bases 1 to 1284)
AUTHORS Gonzalez, F.A.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1992) Fernando A Gonzalez, Biochemistry and
Molecular Biology, University of Massachusetts Medical School, 373
Plantation St., Worcester, MA, 01605, USA

COMMENT On May 6, 2005 this sequence version replaced gi:250709.
FEATURES
Location/Qualifiers
1..1284

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pBluescript II"
/cell_line="Hela"
/clone_lib="Lambda Zap II"
1..1068
CDS

		/organism="synthetic construct"			
		/mol_type="mRNA"			
		/db_xref="taxon:32630"			
		/clone="PLH135362.01X"			
		/lab_host="Escherichia coli DH5alpha T1 resistant"			
		/note="derived from MGC template"			
gene		1. .1140			
		/gene="MAPK3"			
CDS		1. .1140			
		/gene="MAPK3"			
		/codon_start=1			
		/transl_table=11			
		/product="mitogen-activated protein kinase 3"			
		/protein_id="AA342400.1"			
		/db_xref="GI:61363491"			
		/translations="MAAAAGCGGGPRRTGPGVGPGEVMMVKQPPDVGPRVYTG			
		LQVIGSAGVMSADVHKRTVAJKKISPFHOTYCAQTLREIQILLRPHENVIGS			
		IRDLIRASTLEARDVYIVODLMETDLYKLLKQQSLNDHICYFLQILRLGLKYIHS			
		NVLRDLKPSNLINTTCDUKICDFGLARIADPEHDHTGFLTGYVATRWTRAFBIIMLN			
		SKGYTSDITWSYGCILAEMLSNRPYPGKHYLDQNLHILGILGSPSQEDLNCIIMNK			
		ARNYLOSIPSKTVAWAKLPKSDSKALDILLRMLTFNPKNKRTVBEALAHPLYEQY			
		DPTDEPVAEESPFTFAMELDDLPERLKELIFFQETARFPQGVLEAP"			
ORIGIN					
		Query Match	62.1%;	Score 1140;	DB 11; Length 1140;
		Best Local Similarity	100.0%;	Prod. No. 4.3e-190;	
		Matches 1140;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	12	ATGCGCGCGCGCGCGCTCAGGGGGGCGGGGGCGGGAGCCCGTAGAAACGAGGGGGTC	71		
DB					
	1	ATGCGCGCGCGCGCGCTCAGGGGGGCGGGGGCGGGAGCCCGTAGAAACGAGGGGGTC	60		
QY	72	GGCCCGGGGCTCCGGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTTCGACGTGGGCGCG	131		
DB	61	GGCCCGGGGCTCCGGGGGAGGTGGAGATGGTGAAGGGGCGAGCGTTCGACGTGGGCGCG	120		
QY	132	CGGTACACGCAGTTGCAGTACATCGGCGAGGGCGCGTACGGCATGTGTCAAGTTCGGCCTAT	191		
DB	121	CGGTACACGCAGTTGCAGTACATCGGCGAGGGCGCGTACGGCATGTGTCAAGTTCGGCCTAT	180		
QY	192	GACCACTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAACATCAGACC	251		
DB	181	GACCACTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAACATCAGACC	240		
QY	252	TACTGCCAGCGCAGCTCCGGGAGATCCAGATCCTGCTGGCGTTCGCCCATGAGAATGTC	311		
DB	241	TACTGCCAGCGCAGCTCCGGGAGATCCAGATCCTGCTGGCGTTCGCCCATGAGAATGTC	300		
QY	312	ATCGGCATCCGAGACATTTCTCGGGCGTCCACCTCGAAGCCATGAGAGATGTCTACATT	371		
DB	301	ATCGGCATCCGAGACATTTCTCGGGCGTCCACCTCGAAGCCATGAGAGATGTCTACATT	360		
QY	372	GTGCAGACCTGATGGAGACTGACCTGTACAGTTGCTGAAGACCGCAGCAGCTGAGCAAT	431		
DB	361	GTGCAGACCTGATGGAGACTGACCTGTACAGTTGCTGAAGACCGCAGCAGCTGAGCAAT	420		
QY	432	GACCATATCTGTAATTCTCTACAGATCCTCGGGGCGCTCAAAGTACATCCACTCGGCC	491		
DB	421	GACCATATCTGTAATTCTCTACAGATCCTCGGGGCGCTCAAAGTACATCCACTCGGCC	480		
QY	492	AACTGTCTCCACCGAGATCTAAAGCCCTCCAACTGTCTCAATCAACCACTCGGACTTC	551		
DB	481	AACTGTCTCCACCGAGATCTAAAGCCCTCCAACTGTCTCAATCAACCACTCGGACTTC	540		
QY	552	AAGATTGTGATTTTCGGCTGGCCCGGATTCGCCATCCTGAGCATGACCAACACCGGCTTC	611		
DB	541	AAGATTGTGATTTTCGGCTGGCCCGGATTCGCCATCCTGAGCATGACCAACACCGGCTTC	600		
QY	612	CTGACGGAGTATGTGGCTTACCGCTGGTGTACCGGGCCCGCAGAGATCATGCTGAACTCC	671		
DB	601	CTGACGGAGTATGTGGCTTACCGCTGGTGTACCGGGCCCGCAGAGATCATGCTGAACTCC	660		
QY	672	GGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCATTTCTGGGTGAGATGCTCTCT	731		

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:22:44 ; Search time 1117.96 Seconds
(without alignments)
10951.209 Million cell updates/sec

Title: US-10-623-108-5

Perfect score: 1837

Sequence: 1 gagagtgagatggcgcg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2010s.*

5: Geneseqn2020s.*

6: Geneseqn2030s.*

7: Geneseqn2040s.*

8: Geneseqn2050s.*

9: Geneseqn2060s.*

10: Geneseqn2070s.*

11: Geneseqn2080s.*

12: Geneseqn2090s.*

13: Geneseqn2100s.*

14: Geneseqn2110s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1837	100.0	1837	14	ADW12904 Human SMA
2	1725	93.9	1813	13	ACF87406 Human SIR
3	1708.4	93.0	1866	2	AAV62478 Human MAP
4	1708.4	93.0	1866	6	ABK84604 Human CDN
5	1708.4	93.0	1866	8	ACA89915 Gene diff
6	1708.4	93.0	1866	12	ADO24424 Human PRO
7	1708.4	93.0	1866	14	ADZ10042 Human bre
8	1708.4	93.0	3732	6	ABK90802 DNA encod
9	1707	92.9	1777	14	ADW12906 Human SMA
10	1694	92.2	1868	4	AAI59696 Human pol
11	1679.8	91.4	1869	4	AAI57910 Human pol
12	1595	86.8	1726	14	ADW12902 Human SMA
13	1451	79.0	1654	14	ADW12900 Human SMA
14	1437	78.2	1699	12	ADH48413 Human KPP
15	1138.4	62.0	1140	10	ADH59631 Erkl1 enco
16	1138.4	62.0	1896	2	AAV71022 Erkl1-gree
17	1138.4	62.0	1896	3	ADAO0036 EGFP-Erkl1
18	1128	61.4	1305	13	ACN41199 Human dia
19	1098	59.8	1290	13	ACN41200 Human dia

20	1080.2	58.8	1125	10	ABZ77131 Human pro
21	1076	58.6	1214	3	AAC99095 Human pan
22	1066.8	58.1	1071	2	ABA96212 Extracell
23	1061	57.8	1783	14	ADX26048 Novel cel
24	1043.2	56.8	1147	8	ABX72224 Human NOV
25	1021.8	55.6	1747	2	AAQ20260 ERK1. 8/2
26	1016	55.3	1254	13	ACN41201 Human dia
27	986	53.7	1239	13	ACN41202 Human dia
28	965.4	52.6	1238	13	ADV41142 Rat cardl
29	965.4	52.6	1238	14	ADX26182 Novel cel
30	958.6	52.2	1875	2	AAV71030 Erkl1-gree
31	928.4	50.5	1210	13	ACN41203 Human dia
32	851.4	35.5	1484	12	ADG86026 Human tum
33	649.8	35.4	2934	13	ADP56267 Human PRO
34	649.8	35.4	2934	13	ADG92937 Mitogen-a
35	649.8	35.4	2934	14	ADY15945 DNA encod
36	646.6	35.2	1611	2	AAV62479 Human MAP
37	646.6	35.2	1611	6	ABL59534 Human mit
38	646.6	35.2	1611	6	ABK90803 DNA encod
39	646.6	35.2	1611	8	ACFI2891 Human cer
40	646.6	35.2	1611	14	ADY20280 DNA encod
41	646	35.2	3409	10	ADF81745 Leukaemia
42	642	34.9	2791	6	ABK84759 Human CDN
43	642	34.9	2791	10	ACA56665 Human sig
44	642	34.9	2791	12	ADI56461 Human pol
45	641.8	34.9	1083	14	ADV42655 Human psy

ALIGNMENTS

RESULT 1
ADW12904
ID ADW12904 standard; cDNA; 1837 BP.
AC ADW12904;
XX
XX 07-APR-2005 (first entry)
XX Human SMAPK3V3 variant gene.
XX
XX DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
XX Gene; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 12..1151
XX /*tag= a
XX /*product= "SMAPK3V3 protein"
XX
XX US2005013817-A1.
XX
XX 20-JAN-2005.
XX
XX 18-JUL-2003; 2003US-00623108.
XX
XX 18-JUL-2003; 2003US-00623108.
XX
XX (DAIK/) DAI K.
XX Dai K;
XX WPI; 2005-080923/09.
XX P-PSDB; ADW12905.
XX
XX New isolated SMAPK3 polypeptides and nucleic acids, useful for diagnosing
XX diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene
XX in a mammal.
XX Claim 4; SEQ ID NO 5; 55pp; English.
XX The invention relates to novel isolated polypeptide (I) comprising an

CC amino acid sequence selected from sequences comprising 335 or 359 amino acids (ADW12901 or ADW12903) or its fragments. The polypeptides and polynucleotides are useful for diagnosing diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene in a mammal. The fragments of the polypeptides and polynucleotides can also be used as primers or probes. This sequence corresponds to the cDNA encoding the SMAPK3V1 variant protein.

XX
SQ Sequence 1837 BP; 369 A; 593 C; 526 G; 349 T; 0 U; 0 Other;

Query Match 100.0%; Score 1837; DB 14; Length 1837;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGCGGGAGCCCCGTAGAA 60
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 GAGGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGCGGGAGCCCCGTAGAA 60
QY 61 CCAGAGGGGTGGCGCGCGGGGTCGCCGGGAGGTGGAGATGGTGAAGGGGCGAGCGGTTG 120
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 CCAGAGGGGTGGCGCGCGGGGTCGCCGGGAGGTGGAGATGGTGAAGGGGCGAGCGGTTG 120
QY 121 ACCTGGGCGCGCTACACGAGTTGCGAGTACATCGGCGAGGCGGCTACGGCATGGTCA 180
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 ACCTGGGCGCGCTACACGAGTTGCGAGTACATCGGCGAGGCGGCTACGGCATGGTCA 180
QY 181 GCTCGGCTATGACAGCTGGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCTTCG 240
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 GCTCGGCTATGACAGCTGGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCTTCG 240
QY 241 AACATCAGACCTACTCGAGCGACGCTCGGGAGATCCAGATCTCGCTGCGCTCCGCG 300
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 AACATCAGACCTACTCGAGCGACGCTCGGGAGATCCAGATCTCGCTGCGCTCCGCG 300
QY 301 ATGAGAAATGTCATCGGCATCCGAGACATCTCGGGCGTCCACCTCGGAAGCCATGAGAG 360
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 ATGAGAAATGTCATCGGCATCCGAGACATCTCGGGCGTCCACCTCGGAAGCCATGAGAG 360
QY 361 ATGTCTACATTTGCGAGGACTGATGGAGACTGACCTGTACAAAGTTGCTGAAAGCCAGC 420
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 ATGTCTACATTTGCGAGGACTGATGGAGACTGACCTGTACAAAGTTGCTGAAAGCCAGC 420
QY 421 AGCTGAGCAATGACCAATATCTGCTACTTCTCTACAGATCTCTCGGGGCTCAAGTACA 480
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
421 AGCTGAGCAATGACCAATATCTGCTACTTCTCTACAGATCTCTCGGGGCTCAAGTACA 480
QY 481 TCCACTCCGCAACGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 TCCACTCCGCAACGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
QY 541 CCTGCGACCTTAAGATTTGATTTTGGGCTGGCCCGGATGCGGATCCTGAGCATGACC 600
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 CCTGCGACCTTAAGATTTGATTTTGGGCTGGCCCGGATGCGGATCCTGAGCATGACC 600
QY 601 ACACCGGCTTCGAGCGGATGCTGCTACGCGCTGCTACCGGGCCCGAGAGATCATGC 660
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
601 ACACCGGCTTCGAGCGGATGCTGCTACGCGCTGCTACCGGGCCCGAGAGATCATGC 660
QY 661 TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTCGGCTG 720
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
661 TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTCGGCTG 720
QY 721 AGATGCTCTTAACCGGCGCATCTTCCCTGGGCAAGCACTACCTGGATAGCTCAACACCA 780
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
721 AGATGCTCTTAACCGGCGCATCTTCCCTGGGCAAGCACTACCTGGATAGCTCAACACCA 780
QY 781 TTCTGGGCACTCCTGGGCTCCCATCCAGAGGACCTGAATTCATCAACATCAAGG 840
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
781 TTCTGGGCACTCCTGGGCTCCCATCCAGAGGACCTGAATTCATCAACATCAAGG 840
QY 841 CCCGAAACTACCTACAGTCTCTGCCCTCCAAAGCAGGAGTGGCTTGGGCCAAGCTTTCC 900
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
841 CCCGAAACTACCTACAGTCTCTGCCCTCCAAAGCAGGAGTGGCTTGGGCCAAGCTTTCC 900

RESULT 2
ACF87406
ID ACF87406 standard; DNA; 1813 BP.
XX ACF87406;
XX AC
DT 02-JUN-2005 (first entry)

QY 901 CCAAGTCCAGACTCCAAAGCCCTTGACCTGCTGACCGGATGTTAACTTTAACCCCAATA 960
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
901 CCAAGTCCAGACTCCAAAGCCCTTGACCTGCTGACCGGATGTTAACTTTAACCCCAATA 960
QY 961 AACCGATCACAGTGGAGGAAGCGCTGGCTCAACCCCTACTCGGAGCAGTACTATGACCCGA 1020
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
961 AACCGATCACAGTGGAGGAAGCGCTGGCTCAACCCCTACTCGGAGCAGTACTATGACCCGA 1020
QY 1021 CGGATGAGCAGTGGCCGAGGAGCCCTTCACTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1080
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1021 CGGATGAGCAGTGGCCGAGGAGCCCTTCACTTCCAGGAGTGTGGATGAGTACCTTACCTTA 1080
QY 1081 AGGAGCGGCTGAAGAGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1140
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1081 AGGAGCGGCTGAAGAGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1140
QY 1141 AGGCCCCCTAGCCCCAGACAGACATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAG 1200
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1141 AGGCCCCCTAGCCCCAGACAGACATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAG 1200
QY 1201 AGGCAAGAGTCACTGAGGGCTCTGTCAACCCAGGACCTGCTGCTGCTGCTGCTGCTGCT 1260
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1201 AGGCAAGAGTCACTGAGGGCTCTGTCAACCCAGGACCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 CGCCAGACTGTTAGAAATGGACACTGTGCCAGCCCGAGACCTTGGCAGCCCGAGCCGGG 1320
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1261 CGCCAGACTGTTAGAAATGGACACTGTGCCAGCCCGAGACCTTGGCAGCCCGAGCCGGG 1320
QY 1321 GTGGAGCATGGGCTTGGCCACCTCTCTCTTTGTGAGGCTTCCAGCTTCAGGAGGCCA 1380
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1321 GTGGAGCATGGGCTTGGCCACCTCTCTCTTTGCTGAGGCTTCCAGCTTCAGGAGGCCA 1380
QY 1381 AGGCTTCTCTCCCAACCCGCTTCCCAAGGGGCTCGGAGCTCAGTGTGGCCCCAGT 1440
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1381 AGGCTTCTCTCTCCCAACCCGCTTCCCAAGGGGCTCGGAGCTCAGTGTGGCCCCAGT 1440
QY 1441 TCAATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1441 TCAATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1501 TCTGGAATGGAAGGTTCTGGCTGCCCAACCTGTGGAAGGCGAGAGTGGAGGTGGGG 1560
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1501 TCTGGAATGGAAGGTTCTGGCTGCCCAACCTGTGGAAGGCGAGAGTGGAGGTGGGG 1560
QY 1561 GCGCTGAGTAGGAGTCAAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1561 GCGCTGAGTAGGAGTCAAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1621 AGTTTCCCTGGAAGGAACATTCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCCG 1680
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1621 AGTTTCCCTGGAAGGAACATTCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCCG 1680
QY 1681 GGGCGAATCCCTCCCTGCTCAAAAGCTGCTCACTTCCGCTGCTGCTGCTGCTGCTGCTG 1740
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1681 GGGCGAATCCCTCCCTGCTCAAAAGCTGCTCACTTCCGCTGCTGCTGCTGCTGCTGCTG 1740
QY 1741 GTGAGCAGAAGTGGAGTGGGGGGGCGTGAGAGCCCGGGGCCCCCTGCCACCTCCCTGACC 1800
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1741 GTGAGCAGAAGTGGAGTGGGGGGGCGTGAGAGCCCGGGGCCCCCTGCCACCTCCCTGACC 1800
QY 1801 CGTCTAATATATAAATATAGAGATGTGTCTATGGCTG 1837
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1801 CGTCTAATATATAAATATAGAGATGTGTCTATGGCTG 1837

QY 1321 GTGAGCATGGCGCTGGCCACCTCTCTCTTTCTGAGGCTCCAGCTTCAGGACGGCA 1380
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1270 GTGGAGCATGGCGCTGGCCACCTCTCTCTTTCTGAGGCTCCAGCTTCAGGACGGCA 1329
QY 1381 AGGCTTTCTCTCCACCGCCCTCCCAACGGGGCTCGGGAGCTCAGGTGGCCCCAGT 1440
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1330 AGGCTTTCTCTCCACCGCCCTCCCAACGGGGCTCGGGAGCTCAGGTGGCCCCAGT 1389
QY 1441 TCAATCTCCGCTGTCTGTCTGGCGCTTACCTTCCAGGTCCAGTCTCTGGCAGT 1500
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1390 TCAATCTCCGCTGTCTGTCTGGCGCTTACCTTCCAGGTCCAGTCTCTGGCAGT 1449
QY 1501 TCTGGAATGAAGGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGG 1560
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1450 TCTGGAATGAAGGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGG 1509
QY 1561 GCGCTGTAGTGGAGCTCAGGGCCATGCTGCCCCCTCATCTCAATTCATAAACCACCT 1620
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1510 GCGCTGTAGTGGAGCTCAGGGCCATGCTGCCCCCTCATCTCAATTCATAAACCACCT 1569
QY 1621 AGTTTCCCTGAAGAACTTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1680
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1570 AGTTTCCCTGAAGAACTTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC 1629
QY 1681 GCGCGAATCCCTCCCTGTCAAGCTGTCACTTCGCGTGCCTCGCTCTCTGTGTGTG 1740
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1630 GCGCGAATCCCTCCCTGTCAAGCTGTCACTTCGCGTGCCTCGCTCTCTGTGTGTG 1689
QY 1741 GTGAGCAGAAGTGGAGTGGGGGGCGTGGAGAGCCCGCGGCCCTTCCACCTCCCTGACC 1800
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1690 GTGAGCAGAAGTGGAGTGGGGGGCGTGGAGAGCCCGCGGCCCTTCCACCTCCCTGACC 1749
QY 1801 CGTCTAATATATAAATAGAGATGTGTCTATGGCTG 1837
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1750 CGTCTAATATATAAATAGAGATGTGTCTATGGCTG 1786

RESULT 3
AAV62478
ID AAV62478 standard; cdna; 1866 BP.
XX
AC AAV62478;
XX
DT 18-JAN-1999 (first entry)
XX
DE Human MAP kinase, ERK-1 cDNA sequence.
XX
KW ERK-1; ERK-2; mitogen-activated protein kinase; MAP kinase; human;
KW inhibition; malignant; neoplastic growth; epithelial cell; mammal;
KW endothelial cell; antisense oligonucleotide; primary cancer;
KW metastatic cancer; breast cancer; prostate cancer; angiosarcoma;
KW endocrine tissue cancer; ds.
XX
OS Homo sapiens.
XX
PN W09844101-A1.
XX
PD 08-OCT-1998.
XX
PF 19-MAR-1998; 98WO-US005471.
XX
PR 28-MAR-1997; 97US-00827520.
PR 01-APR-1997; 97US-00831994.
PR 12-AUG-1997; 97US-00909742.
XX
PA (UJNY) UNIV NEW YORK STATE RES FOUND.
XX
PI Sivaraman VS, Wang H, Malbon CC;
XX
DR WPI; 1998-557109/47.
XX
PT Treatment of e.g. breast or prostate cancer or angiosarcoma - by
PT administering antisense oligonucleotides to genes encoding mitogen-

PT activating protein kinases ERK1 and ERK2.
XX
PS Disclosure; Page 11-12; 59pp; English.
XX
CC This represents the cDNA sequence of the human ERK-1, a mitogen-activated
CC protein (MAP) kinase. The invention provides a method of inhibiting
CC malignant neoplastic growth of epithelial or endothelial cell in a mammal
CC which comprises administering to the mammal an effective amount of an
CC oligonucleotide complementary to part of the mRNA for the MAP kinases,
CC ERK-1 or ERK2 which is over-expressed in the mammal. Also provided is a
CC method for identifying and monitoring potentially malignant neoplastic
CC cells by measuring the levels of ERK1 and ERK2 mRNA in epithelial or
CC endothelial cells and comparing it to the levels from normal cells of the
CC same origin. Administration of the ERK1 and ERK2 antisense
CC oligonucleotides (AAV62480 and AAV62481) to neoplastic endothelial or
CC epithelial cells inhibits over-expression of ERK1 and ERK2. This can be
CC used to treat epithelial and endothelial malignancies including primary
CC or metastatic cancers of e.g. the breast, prostate, other endocrine
CC tissue or angiosarcoma
XX
SQ Sequence 1866 BP; 380 A; 605 C; 535 G; 346 T; 0 U; 0 Other;

Query Match 93.0%; Score 1708.4; DB 2; Length 1866;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;

QY 3 GGAGTGGAGATGGCGCGCGCGCTCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 62
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
64 GGAGTGGAGATGGCGCGCGCGCTCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 123
QY 63 GAGGGGGTTCGGGCCCGGGGGTCCCGGGGAGGTGGAGATGTGAAGGGGCGAGCCGTTCGAC 122
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
124 GAGGGGGTTCGGGCCCGGGGGTCCCGGGGAGGTGGAGATGTGAAGGGGCGAGCCGTTCGAC 183
QY 123 GTGGGCGCGCGCTACACGCACTTTCAGTACATCTGGGGGGGGGGGGGGGGGGGGGG 182
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
184 GTGGGCGCGCGCTACACGCACTTTCAGTACATCTGGGGGGGGGGGGGGGGGGGGGG 243
QY 183 TCGGCTATGACACGCTGGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAA 242
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
244 TCGGCTATGACACGCTGGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCGAA 303
QY 243 CATCAGACCTACTGCGCAGCGCACGCTCCGGGAGATCCAGATCCTGCTGCGCTTCGCGCAT 302
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
304 CATCAGACCTACTGCGCAGCGCACGCTCCGGGAGATCCAGATCCTGCTGCGCTTCGCGCAT 363
QY 303 GAGATGTTCATCGGCATTCGAGACATTTTCGGGGGTCCACCTGGAAGCCATGAGAGAT 362
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
364 GAGATGTTCATCGGCATTCGAGACATTTTCGGGGGTCCACCTGGAAGCCATGAGAGAT 423
QY 363 GTCTACATTTGCAGGACCTGATGGAGACTGACCTGTACAACTTGTCTGAAAAGCCAGCAG 422
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
424 GTCTACATTTGCAGGACCTGATGGAGACTGACCTGTACAACTTGTCTGAAAAGCCAGCAG 483
QY 423 CTGAGCAATGACCATATCTGTACTTCTCTACAGATCTCTCGGGGGCTCAAGTACATC 482
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
484 CTGAGCAATGACCATATCTGTACTTCTCTACAGATCTCTCGGGGGCTCAAGTACATC 543
QY 483 CATCTCGCAACGCTGCTCCAGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACC 542
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
544 CATCTCGCAACGCTGCTCCAGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACC 603
QY 543 TCGGACCTTAAGATTTGTGATTTTCGGCCCTGGCCGATTTGGCGATTCCTGAGCATGACCA 602
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
604 TCGGACCTTAAGATTTGTGATTTTCGGCCCTGGCCGATTTGGCGATTCCTGAGCATGACCA 663
QY 603 ACCGCTTCTCTGACGAGTATGTGGCTACGCGCTGTGTACCGGGCCCCCAGAGATCATGTG 662
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
664 ACOGGCTTCTGACGAGTATGTGGCTACGCGCTGTGTACCGGGCCCCCAGAGATCATGTG 723
QY 663 AACTCAAGGGCTATACCAAGTCCATCGCATCTGTGTGGGTGCGATCTTGGGTGAG 722
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
724 AACTCAAGGGCTATACCAAGTCCATCGCATCTGTGTGGGTGCGATCTTGGGTGAG 783

QY 723 ATGCTCTCTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCCACATT 782
Db 784 ATGCTCTCTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCCACATT 843
QY 783 CTGGGCACTCTGGGCTCCCAATCCAGGAGACCTGAATTGTATCATCAATGAAGGCC 842
Db 844 CTGGGCACTCTGGGCTCCCAATCCAGGAGACCTGAATTGTATCATCAATGAAGGCC 903
QY 843 CGAACTACCTACAGTCTCTGCGCTCCCAAGACCAAGGTGGTGGGCAAGCTTTTCCCC 902
Db 904 CGAACTACCTACAGTCTCTGCGCTCCCAAGACCAAGGTGGTGGGCAAGCTTTTCCCC 963
QY 903 AAGTCAGACTCCAAAGCGCTTGACCTGCTGACCGGATGTTAACTTTAAACCCCAATAAA 962
Db 964 AAGTCAGACTCCAAAGCGCTTGACCTGCTGACCGGATGTTAACTTTAAACCCCAATAAA 1023
QY 963 CGGATCAGTGGAGAGCGCTGGCTCAACCCCTACCTGAGCAGTACTATGACCCGACG 1022
Db 1024 CGGATCAGTGGAGAGCGCTGGCTCAACCCCTACCTGAGCAGTACTATGACCCGACG 1083
QY 1023 GATGAGCCAGTGGCGGAGGAGCGCTTCACTTTCGCCATGGAGCTGGATGACCTACCTAAG 1082
Db 1084 GATGAGCCAGTGGCGGAGGAGCGCTTCACTTTCGCCATGGAGCTGGATGACCTACCTAAG 1143
QY 1083 GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTTCAGCCCGGAGTCTGGAG 1142
Db 1144 GAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTTCAGCCCGGAGTCTGGAG 1203
QY 1143 GCGCCCTAGCCAGACAGACATCTCTGCACTGGGCGCTGGAAACAGAAAGAG 1202
Db 1204 GCGCCCTAGCCAGACAGACATCTCTGCACTGGGCGCTGGAAACAGAAAGAG 1243
QY 1203 GCAAGAGTCACTGAGGCGCTCTGTCAACAGGAGCTGCTCTGCGCTGCGCTCTCCCG 1262
Db 1244 -----GGACCTGCGCTCTGCGCTGCGCTCTCTCCCG 1272
QY 1263 CCAGACTGTTAGAAAATGGACACTGTGCCAGCGCGGACCTTGGCAGCCAGGCGGGGT 1322
Db 1273 CCAGACTGTTAGAAAATGGACACTGTGCCAGCGCGGACCTTGGCAGCCAGGCGGGGT 1332
QY 1323 GGAGCATGGGCTGGCCACTCTCTCTCTTTGCTGAGGCTTCAGGCTTCAGGAGCCAG 1382
Db 1333 GGAGCATGGGCTGGCCACTCTCTCTCTTTGCTGAGGCTTCAGGCTTCAGGAGCCAG 1392
QY 1383 GCCTTCTCTCCACCGCGCTCCCAAGCGGCTCGGAGCTCAGTGSCCCAGTTC 1442
Db 1393 GCCTTCTCTCCACCGCGCTCCCAAGCGGCTCGGAGCTCAGTGSCCCAGTTC 1452
QY 1443 AATCTCCC ---GCTGCTGCTGCTGCGCTTTACCTTCCAGCGCTCCAGTCTCTGGCAG 1499
Db 1453 AATCTCCGCTGCTGCTGCTGCTGCGCTTTACCTTCCAGCGCTCCAGTCTCTGGCAG 1512
QY 1500 TTCTGGAATGAAGGTTCTGGCTGCCCAACCTGCTGAAGGCGAGAGTGGAGGGTGG 1559
Db 1513 TTCTGGAATGAAGGTTCTGGCTGCCCAACCTGCTGAAGGCGAGAGTGGAGGGTGG 1572
QY 1560 GGGCGCTAGTAGGACTCAGGCGCATGCTGCGCCCTCTCATCTCATTAACCCACCC 1619
Db 1573 GGGCGCTAGTAGGACTCAGGCGCATGCTGCGCCCTCTCATCTCATTAACCCACCC 1632
QY 1620 TAGTTTCTCTGAAGAACATTCTTAGTCTCAAGGGCTAGCATCTCTGAGGAGCCAGGCC 1679
Db 1633 TAGTTTCTCTGAAGAACATTCTTAGTCTCAAGGGCTAGCATCTCTGAGGAGCCAGGCC 1692
QY 1680 GGGCGGAATCCCTCCCTCTCAAGCTGTCACTTCGCTGCGCTGCTGCTTCTGTGT 1739
Db 1693 GGGCGGAATCCCTCCCTCTCAAGCTGTCACTTCGCTGCGCTGCTGCTTCTGTGT 1752
QY 1740 GGTGAGCAGAGTGAAGCTGGGGGCGCTGGAGAGCCCGCGCCCTGCGACCTCCCTGAC 1799
Db 1753 GGTGAGCAGAGTGAAGCTGGGGGCGCTGGAGAGCCCGCGCCCTGCGACCTCCCTGAC 1812

QY 1800 CCGTCTAATATATAAATATAGAGATGCTCTATGGCTG 1837
Db 1813 CCGTCTAATATATAAATATAGAGATGCTCTATGGCTG 1850

RESULT 4

ABK84604
ID ABK84604 standard; cDNA; 1866 BP.
XX ABK84604;
AC
XX
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #1175.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.

XX WO200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US030821.

XX 03-OCT-2000; 2000US-0237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 1175; 114pp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's

CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1866 BP; 380 A; 605 C; 535 G; 346 T; 0 U; 0 Other;

Query Match 93.0%; Score 1708.4; DB 6; Length 1866;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;

QY 3 GGAGTGGAGATGCGCGCGCGCGCGGTCTAGGGGGGCGGGGCGGGAGCCCGTAGAAC 62
DB 64 GGAGTGGAGATGCGCGCGCGCGGTCTAGGGGGGCGGGGCGGGAGCCCGTAGAAC 123

QY 63 GAGGGGTGCGCGCGCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTTCGAC 122
DB 124 GAGGGGTGCGCGCGCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTTCGAC 183

QY 123 GTGGGCGCGGCTACACGCAAGTTGCAGTATCATCGCGAGGGCGGTACGGCATGGTTCAGC 182
DB 184 GTGGGCGCGGCTACACGCAAGTTGCAGTATCATCGCGAGGGCGGTACGGCATGGTTCAGC 243

QY 183 TCGGGCTATGACCAACGTCGCGCAAGACTGCGTGGCCATCAAGAAGATCAGCCCTTCGAA 242
DB 244 TCGGCTATGACCAACGTCGCGCAAGACTGCGTGGCCATCAAGAAGATCAGCCCTTCGAA 303

QY 243 CATCAGACCTACTGCCAGCGCAGCTCCGGAGATCCAGATCCGTGCGCTCCGCCAT 302
DB 304 CATCAGACCTACTGCCAGCGCAGCTCCGGAGATCCAGATCCGTGCGCTCCGCCAT 363

QY 303 GAGATGTCATCGGCATCCGAGACATTCGCGGGCTCCACCCTGGAAGCATGAGAGAT 362
DB 364 GAGATGTCATCGGCATCCGAGACATTCGCGGGCTCCACCCTGGAAGCATGAGAGAT 423

QY 363 GTCTAATTGTGAGGACCTGATGGAGACTGACCTGTAAGTTGCTGAAAGCCAGAC 483
DB 424 GTCTAATTGTGAGGACCTGATGGAGACTGACCTGTAAGTTGCTGAAAGCCAGAC 483

QY 423 CTGAGCAATGACCATCTGCTACTTCTTACAGATCTCGGGGCTCCAGGTACATC 482
DB 484 CTGAGCAATGACCATCTGCTACTTCTTACAGATCTCGGGGCTCCAGGTACATC 543

QY 483 CACTCCGCGCAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACC 542
DB 544 CACTCCGCGCAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCAGCAACACC 603

QY 543 TGGCACTTAAAGATTTGTGATTTGCGCTCGGCCGGATTCGCCATCTGAGCATGACAC 602
DB 604 TGGCACTTAAAGATTTGTGATTTGCGCTCGGCCGGATTCGCCATCTGAGCATGACAC 663

QY 603 ACCGGCTTCTGACGGAGTATGCGTACGCTGTTACCGGGCCCGCAGAGATCATGCTG 662
DB 664 ACCGGCTTCTGACGGAGTATGCGTACGCTGTTACCGGGCCCGCAGAGATCATGCTG 723

QY 663 AACTCCAAAGGGCTATACCAAGTCCATCGCATCTGCTGCTGCGGCTGCAATTCGGCTGAG 722
DB 724 AACTCCAAAGGGCTATACCAAGTCCATCGCATCTGCTGCTGCGGCTGCAATTCGGCTGAG 783

QY 723 ATGCTCTTAAACCGGCCATCTTCCCTGGCAAGCACTTACTGGATCAGCTCAACCAATT 782
DB 784 ATGCTCTTAAACCGGCCATCTTCCCTGGCAAGCACTTACTGGATCAGCTCAACCAATT 843

QY 783 CTGGGCATCTGGGCTCCCATCCGAGGAGCTGAATTTGATCATCAACATGAAGGCC 842
DB 844 CTGGGCATCTGGGCTCCCATCCGAGGAGCTGAATTTGATCATCAACATGAAGGCC 903

QY 843 CGAAACTACATGCTCTGCGCTCCCAAGCAAGGTGGCTTGGGCGAAGCTTTTCCCC 902

DB 904 CGAAACTACCTACAGTCTCTGCGCTCCCAAGACCAAGTGGCTTGGGCCCAAGCTTTTCCCC 963
QY 903 AAGTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAA 962
DB 964 AAGTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAA 1023
QY 963 CGGATCACAGTGGAGGAAGCGTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG 1022
DB 1024 CGGATCACAGTGGAGGAAGCGTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG 1083

QY 1023 GATGACCATGTCGCGAGAGGCCCTTACCTTGGCATGGAGTGGAGTGGATGACCTACTAAG 1082
DB 1084 GATGACCATGTCGCGAGAGGCCCTTACCTTGGCATGGAGTGGAGTGGATGACCTACTAAG 1143

QY 1083 GAGCGGCTGAAGAGGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1142
DB 1144 GAGCGGCTGAAGAGGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGGAG 1203

QY 1143 GCCCCTTAGCCCGACAGACATCTCTGCAACCCTGGGGCCTGGAAACAGAACTGGCAAAGAG 1202
DB 1204 GCCCCTTAGCCCGACAGACATCTCTGCAACCCTGGGGCCT----- 1243

QY 1203 GCAAGAGTCACTGAGGGCCTCTGTACCCAGCACTGTGCTCTGCTGCTGCCCTCTCCCG 1262
DB 1244 -----GGACCTGCTCTGCTGCCCTCTCCCG 1272

QY 1263 CCAGACTGTTAGAAATGAGACACTGTCGCCAGCCGAGCTTGGCAGCCAGCGGGGT 1322
DB 1273 CCAGACTGTTAGAAATGAGACACTGTCGCCAGCCGAGCTTGGCAGCCAGCGGGGT 1332

QY 1323 GGAGCATGGGCTGCGCACCTCTCTTCTTGTGAGGGCTCCAGCTTTCAGGCAGGCGCAAG 1382
DB 1333 GGAGCATGGGCTGCGCACCTCTCTCTTGTGAGGGCTCCAGCTTTCAGGCAGGCGCAAG 1392

QY 1383 GCTTCTCTCCCAACCGCCCTTCCCAAGCGGGCTCGGAGCTCAGGTGGGCCCCAGTTTC 1442
DB 1393 GCTTCTCTCCCAACCGCCCTTCCCAAGCGGGCTCGGAGCTCAGGTGGGCCCCAGTTTC 1452

QY 1443 AATCTCCC---GCTGCTGCTGCGGCCCTTACCTTCCCAGCGCTCCAGTCTCTGGCAG 1499
DB 1453 AATCTCCCCTGCTGCTGCGGCCCTTACCTTCCCAGCGCTCCAGTCTCTGGCAG 1512

QY 1500 TTCTGGAATGGAAGGTTCTGGCTGCCCAACCTTCTGTAAGGGCAGAGGTGGAGGGTGGG 1559
DB 1513 TTCTGGAATGGAAGGTTCTGGCTGCCCAACCTTCTGTAAGGGCAGAGGTGGAGGGTGGG 1572

QY 1560 GGGCGTGAATGAGGACTCAGGGCCATGCTGCCGCCCTCATCTCATTTCAAACCCCAACC 1619
DB 1573 GGGCGTGAATGAGGACTCAGGGCCATGCTGCCGCCCTCATCTCATTTCAAACCCCAACC 1632

QY 1620 TAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCC 1679
DB 1633 TAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCC 1692

QY 1680 GGGCGGAATCCCTCTCCCTGTCAAAGTGTCACTTGGCGTGGCCCTGCTGCTTCTGTGTGT 1739
DB 1693 GGGCGGAATCCCTCTCCCTGTCAAAGTGTCACTTGGCGTGGCCCTGCTGCTTCTGTGTGT 1752

QY 1740 GGTGAGCAAGTGAAGTGGGGGGGTGGAGAGCCCGGCGCCCTGCCACCTCCCTGAC 1799
DB 1753 GGTGAGCAAGTGAAGTGGGGGGGTGGAGAGCCCGGCGCCCTGCCACCTCCCTGAC 1812

QY 1800 CCGTCTTAATATATAATATAGAGATGTCTATGCTG 1837
DB 1813 CCGTCTTAATATATAATATAGAGATGTCTATGCTG 1850

RESULT 5
AC89915
ID AC89915 standard; cDNA; 1866 BP.
XX AC
XX AC89915;

Db 544 ||||| CACTCGCCCAACGTGCTCCACCGAGATCTTAAGCCCTCCAACTGCTCAGCAACACACC 603
QY 543 TCGGACCTTAAGATTTGATTTGCGCTCGCCCGGATTTGCCCATCTCGAGCATGACCCAC 602
Db 604 TCGGACCTTAAGATTTGATTTGCGCTCGCCCGGATTTGCCCATCTCGAGCATGACCCAC 663
QY 603 ACCGGCTTCTCGACGAGTATGGCTACGCGTGTGTACCGGGCCCGAGAGATCATGCTG 662
Db 664 ACCGGCTTCTCGACGAGTATGGCTACGCGTGTGTACCGGGCCCGAGAGATCATGCTG 723
QY 663 AACTCAAGGGCTATACCAAGTCTCATGATCTGCTGTGCTGTGCTGTGCTGTGCTGTG 722
Db 724 AACTCAAGGGCTATACCAAGTCTCATGATCTGCTGTGCTGTGCTGTGCTGTGCTGTG 783
QY 723 ATGCTCTCTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGATCAGCTCAACCAAT 782
Db 784 ATGCTCTCTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGATCAGCTCAACCAAT 843
QY 783 CTGGGCACTCTGGGCTCCCATCCAGGAGACCTGAATTTGATCATCAACATGAAGGCC 842
Db 844 CTGGGCACTCTGGGCTCCCATCCAGGAGACCTGAATTTGATCATCAACATGAAGGCC 903
QY 843 CGAAACTACCTACGTTCTGCTGCTCCCAAGCAAGTGCTTGGCGCAAGCTTTTCCCC 902
Db 904 CGAAACTACCTACGTTCTGCTGCTCCCAAGCAAGTGCTTGGCGCAAGCTTTTCCCC 963
QY 903 AAGTCAGACTCCAAAGCCCTTGACCTGCTGACCGGATTTAACTTTAAACCCCAATAA 962
Db 964 AAGTCAGACTCCAAAGCCCTTGACCTGCTGACCGGATTTAACTTTAAACCCCAATAA 1023
QY 963 CGGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG 1022
Db 1024 CGGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG 1083
QY 1023 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTGTGCCATGAGAGTGGATGACCTACCTAAG 1082
Db 1084 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTGTGCCATGAGAGTGGATGACCTACCTAAG 1143
QY 1083 GAGGGCTGAAGGAGCTCATCTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGGAG 1142
Db 1144 GAGGGCTGAAGGAGCTCATCTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGGAG 1203
QY 1143 GCCCCTTAGCCACAGACAGATCTCTGACCCCTGGGCGCTGGAACAGACTGGCAAGAG 1202
Db 1204 GCCCCTTAGCCACAGACAGATCTCTGACCCCTGGGCGCT----- 1243
QY 1203 GCAAGAGGTCACTAGGGCCCTCTGTACCCAGGACCTGCTCTGCTGCTGCTGCTGCTG 1262
Db 1244 -----GGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1272
QY 1263 CCAGACTTTAGAAAATGGACACTGTGCGCCAGCCCGGACCTTGGCAGCCAGGCGGGGT 1322
Db 1273 CCAGACTTTAGAAAATGGACACTGTGCGCCAGCCCGGACCTTGGCAGCCAGGCGGGGT 1332
QY 1323 GGACATGGGCTGCGCCACCTCTCTCTTGTGCTGAGGCTTCCAGCTTCAGGAGGCCAAG 1382
Db 1333 GGACATGGGCTGCGCCACCTCTCTCTTGTGCTGAGGCTTCCAGCTTCAGGAGGCCAAG 1392
QY 1383 GCCTTCTCTCCCAACCGCCCTCCCAACCGGGGCTCGGAGCTTCAGGTGGCCCAATTC 1442
Db 1393 GCCTTCTCTCCCAACCGCCCTCCCAACCGGGGCTCGGAGCTTCAGGTGGCCCAATTC 1452
QY 1443 AATCTCCC---GCTGCTGCTGCTGCTGCTTCCCTGCGAGCTTCCAGCTTCTCTGCGAG 1499
Db 1453 AATCTCCGCTGCTGCTGCTGCTGCTTCCCTGCGAGCTTCCAGCTTCTCTGCGAG 1512
QY 1500 TTCTGGAAATGAAGGTTCTGGCTGCCCACTGCTGAAGGGCAGAGGTGGAGGTGGG 1559
Db 1513 TTCTGGAAATGAAGGTTCTGGCTGCCCACTGCTGAAGGGCAGAGGTGGAGGTGGG 1572
QY 1560 GGGCGCTGAGTGGAGTCAAGGGCATGCTGCCCCCTCATCTCATTTCAACCCCAACC 1619
|||||

Db 1573 GGGCGCTAGTAGGAGACTCAGGGCCATGCTGCCCTCATCTCATTTCAAACCCACCC 1632
QY 1620 TAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCC 1679
Db 1633 TAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGGCC 1692
QY 1680 GGGCCGAATCCCTCCCTGCTCAAGCTGTCACCTTGGGTCGCCCTGCTGCTGCTGCTG 1739
Db 1693 GGGCCGAATCCCTCCCTGCTCAAGCTGTCACCTTGGGTCGCCCTGCTGCTGCTGCTG 1752
QY 1740 GGTGACGAGAAGTGGAGCTGGGGGCTGGAGAGCCCGGCCCTGCCACCTCCCTGAC 1799
Db 1753 GGTGACGAGAAGTGGAGCTGGGGGCTGGAGAGCCCGGCCCTGCCACCTCCCTGAC 1812
QY 1800 CGCTCTAATATATAAATATAGAGATGTCTATGGCTG 1837
Db 1813 CGCTCTAATATATAAATATAGAGATGTCTATGGCTG 1850

RESULT 7
AD210042
ID AD210042 standard; DNA; 1866 BP.
XX
AC AD210042;
XX
DT 30-JUN-2005 (first entry)
XX Human breast cancer marker MAPK3 DNA.
XX neoplasm; tumor marker; cytostatic; breast tumor; gene therapy; da; gene;
XX MAPK3.
XX Homo sapiens.
OS
XX
PN BP1522594-A2.
XX
PD 13-APR-2005.
XX
PF 30-JUN-2004; 2004BP-00015374.
XX
PR 06-OCT-2003; 2003BP-00022587.
XX
XX (PARB) BAYER HEALTHCARE AG.
PA
PI Munnes M, Bojar H;
XX
XX WPI; 2005-265481/28.
DR P-PSDB; AD210062.
XX

Characterizing the state of a neoplastic disease in a subject by
comparing the pattern of determined expression levels of marker genes in
a biological sample with reference patterns of expression levels.
Claim 2; SEQ ID NO 486; 542pp; English.

The invention relates to a novel method for characterizing the state of a
neoplastic disease in a subject comprising determining the pattern of
expression levels of at least 6, 8, 10, 15, 20, 30, 40 or 67 marker genes
in a biological sample, comparing the pattern of expression levels with
reference patterns of expression levels and characterizing the state of
the neoplastic disease in the subject from the outcome of the comparison.
The method of the invention demonstrates cytostatic applications and may
be useful in characterizing the state of a neoplastic disease in a
subject, preferably breast cancer, in gene therapy and for screening for
compounds having a curative effect on a neoplastic disease. The current
sequence is that of the human breast cancer marker MAPK3 DNA of the
invention which is differentially expressed in neoplastic tissue.

Sequence 1866 BP; 380 A; 605 C; 535 G; 346 T; 0 U; 0 Other;
Query Match 93.0%; Score 1708.4; DB 14; Length 1866;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;

QY	3	GGAGTGCAGATGGCGGGCGGGCGGCTCAGGGGGCGGGGCGGGGAGCCCCGTGAGAAC	62
DB	64	GGAGTGCAGATGGCGGGCGGGCGGCTCAGGGGGCGGGGAGCCCCGTGAGAAC	123
QY	63	GAGGGGTGCGCCCGGGGGTCCCGGGGAGGTGTGAGATGTGTGAAGGGGAGCGGTTTCGAC	122
DB	124	GAGGGGTGCGCCCGGGGGTCCCGGGGAGGTGTGAGATGTGTGAAGGGGAGCGGTTTCGAC	183
QY	123	GTGGGCGCGGTACACGCAAGTTCAGTACATCGGCGAGGGCGGTACGATGCTTCGAC	182
DB	184	GTGGGCGCGGTACACGCAAGTTCAGTACATCGGCGAGGGCGGTACGATGCTTCGAC	243
QY	183	TCGGCCCTATGACCAACGTCGCGCAAGACTCGGTGGCCATCAAGAAGATCAGCCCCCTTCGAA	242
DB	244	TCGGCCCTATGACCAACGTCGCGCAAGACTCGGTGGCCATCAAGAAGATCAGCCCCCTTCGAA	303
QY	243	CATCAGACCTACTGCCAGCGCAACGCTCGGGAGATCCAGATCTCTGTGTGGGTTCCGGCCAT	302
DB	304	CATCAGACCTACTGCCAGCGCAACGCTCGGGAGATCCAGATCTCTGTGTGGGTTCCGGCCAT	363
QY	303	GAGAACTCATCGGCATCCGAGACATTTCTGGGGCGTCCACCTTGGAGCCATGAGAGAT	362
DB	364	GAGAACTCATCGGCATCCGAGACATTTCTGGGGCGTCCACCTTGGAGCCATGAGAGAT	423
QY	363	GTCTACATTTGTCAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG	422
DB	424	GTCTACATTTGTCAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGCAG	483
QY	423	CTGAGCAATGACCAATATCTGTACTTCTCTTACCAGATCTCTGGGGGCTTCAAGTACATC	482
DB	484	CTGAGCAATGACCAATATCTGTACTTCTCTTACCAGATCTCTGGGGGCTTCAAGTACATC	543
QY	483	CACTCCGCGAAAGTGTCTCCACCGAGATCTAAAGCCCTCACAACCTGCTATCAACACCAC	542
DB	544	CACTCCGCGAAAGTGTCTCCACCGAGATCTAAAGCCCTCACAACCTGCTATCAACACCAC	603
QY	543	TGGACCTTAAGATTGTGTATTTTCGGCTTGCGCGGATTTGCCGATCTGTAGCATGACCAC	602
DB	604	TGGACCTTAAGATTGTGTATTTTCGGCTTGCGCGGATTTGCCGATCTGTAGCATGACCAC	663
QY	603	ACGGGCTTCTGACGAGATATGTGGCTACGCGCTGTATACGGGGCCCGAGAGATCATGCTG	662
DB	664	ACGGGCTTCTGACGAGATATGTGGCTACGCGCTGTATACGGGGCCCGAGAGATCATGCTG	723
QY	663	AATCTCAAGGGCTATACCAAGTTCATCGACATCTGGTCTGTGGGCTGCATCTTGGCTGAG	722
DB	724	AATCTCAAGGGCTATACCAAGTTCATCGACATCTGGTCTGTGGGCTGCATCTTGGCTGAG	783
QY	723	ATGCTCTCTAACCGGGCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACCATTT	782
DB	784	ATGCTCTCTAACCGGGCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACCATTT	843
QY	783	CTGGGCATCTCTGGGCTCCCCATCCAGAGGACCTGAAATTGTATCATCAATGAAGGCC	842
DB	844	CTGGGCATCTCTGGGCTCCCCATCCAGAGGACCTGAAATTGTATCATCAATGAAGGCC	903
QY	843	CGAAATACCTTACAGTCTCTGCCCTCCAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC	902
DB	904	CGAAATACCTTACAGTCTCTGCCCTCCAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC	963
QY	903	AAGTCAGATCTCAAGCGCTTGACCTGCTGGACCGGATGTTAACTTTTAAACCCCAATAAA	962
DB	964	AAGTCAGATCTCAAGCGCTTGACCTGCTGGACCGGATGTTAACTTTTAAACCCCAATAAA	1023
QY	963	CGGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG	1022
DB	1024	CGGATCACAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACG	1083
QY	1023	GATGAGCCAGTGGCCGAGGAGCCCTTCACTTTCGCCATGAGCTTGGATGACTTACCTTAAG	1082
DB	1084	GATGAGCCAGTGGCCGAGGAGCCCTTCACTTTCGCCATGAGCTTGGATGACTTACCTTAAG	1143

QY	1083	GAGCGGCTGAAGGAGCTCATCTTTCCAGGAGACAGCA CGTTCAGGCCGGGAGTGCTGGAG	1144
Db	1144	GAGCGGCTGAAGGAGCTCATCTTTCCAGGAGACAGCA CGTTCAGGCCGGGAGTGCTGGAG	1203
QY	1143	GCCCCCTAGCCACAGACATCTCTGCACCCCTGGGGCCTTGGAA CAGA A CTGCGAAGAG	1202
Db	1204	GCCCCCTAGCCACAGACATCTCTGCACCCCTGGGGCCT- - - - - GGACCTGCCTCTGCCTCTCTCCCG	1243
QY	1203	GCAAGAGGTCACTGAGGGCCTCTGTCA CC CAGGACCTTGCCTCTCTGCTGCTGCCCTCTCCCG	1262
Db	1244	- - - - - GGACCTGCCTCTGCCTCTCTCCCG	1272
QY	1263	CCAGACTGTTAGAAAATGGA CAT GTGTGCCAGCCCGGACCTTGGCAGCC CAGAGCCGGGGT	1322
Db	1273	CCAGACTGTTAGAAAATGGA CAT GTGTGCCAGCCCGGACCTTGGCAGCC CAGAGCCGGGGT	1332
QY	1323	GGAGCATGGCCCTGGCCACCTCTCTCTTTGTCTGAGGCCCTCAGCTTCAGGCAGGCCAACG	1382
Db	1333	GGAGCATGGCCCTGGCCACCTCTCTCTTTGTCTGAGGCCCTCAGCTTCAGGCAGGCCAACG	1392
QY	1383	GCCTTCTCTCCACCACCGCCCTCCCAC GGGGCGCTTCGGAGCTCAGAGTGGCCCCCAGATTTC	1442
Db	1393	GCCTTCTCTCCACCACCGCCCTCCCAC GGGGCGCTTCGGAGCTCAGAGTGGCCCCCAGATTTC	1452
QY	1443	AATCTCCCC-- - - GCTGTCTGTCTGCGGCCCTTA CTCTCC CAGCGTCCCA GTCTCTGSCAG	1499
Db	1453	AATCTCCCGTGTCTGTCTGTCTGCTGCGGCCCTTAC TTTCC CAGCGTCCCA GTCTCTGSCAG	1512
QY	1500	TTCGTGAATGGAAAGGTTCTGGCTGCGCCCA ACTGTCTGAAGGCAGAGGTGGAGGGTGGG	1559
Db	1513	TTCGTGAATGGAAAGGTTCTGGCTGCGCCCA ACTGTCTGAAGGCAGAGGTGGAGGGTGGG	1572
QY	1560	GGCGCTGAGTGGGACTCAGGGCCATGCTGCC CCCCTCATCTCA TTCAA ACCCCACCC	1619
Db	1573	GGCGCTGAGTGGGACTCAGGGCCATGCTGCC CCCCTCATCTCA TTCAA ACCCCACCC	1632
QY	1620	TAGTTTTCCCTGGAAGGAACA TTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC	1679
Db	1633	TAGTTTTCCCTGGAAGGAACA TTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCC	1692
QY	1680	GGCGCGAATCCCTCCCTGTCAAAGCTGTCACTTCGGGTGCCCTCGCTGCTTCTGTGTGT	1739
Db	1693	GGCGCGAATCCCTCCCTGTCAAAGCTGTCACTTCGGGTGCCCTCGCTGCTTCTGTGTGT	1752
QY	1740	GGTGAGCAGAAGTGGAGCTGGGGGGCGTGGAGAGCCCGCGCCCTGCCACCTCCCTGAC	1799
Db	1753	GGTGAGCAGAAGTGGAGCTGGGGGGCGTGGAGAGCCCGCGCCCTGCCACCTCCCTGAC	1812
QY	1800	CCGCTTAATAATAATAATAGAGATGTCTATGGCTG	1837
Db	1813	CCGCTTAATAATAATAATAGAGATGTCTATGGCTG	1850
RESULT 8			
ABK90802			
ID ID ABK90802 standard; DNA; 3732 BP.			
XX AC ABK90802;			
XX XX 05-NOV-2002 (first entry)			
XX DE DNA encoding human extracellular signal regulated kinase, ERK1.			
XX KW Extracellular signal regulated kinase; hyperalgesia; surgery;			
KW opiod withdrawal; pain sensitisation; analgesic; chronic pain; ERK1;			
KW human; gene; ds.			
XX OS Homo sapiens.			
XX XX			
FH	Key	Location/Qualifiers	
FT	CDS	73..1212	
FT		/*tag= a	
FT		/product= "ERK1"	

/note= "Extracellular signal regulated kinase"

[illegible]

Db 970 -----CAGTGGCCGAGAGCCCTTCACTTGGCCATGGAGCTGGATGACCTACCTA 1020
QY 1081 AGGAGCGCTGAAGAGCTCATCTTCCAGGAGACAGACACGCTTCCAGCCCGAGTGTGG 1140
Db 1021 AGGAGCGCTGAAGAGCTCATCTTCCAGGAGACAGACACGCTTCCAGCCCGAGTGTGG 1080
QY 1141 AGGCCCCCTAGCCAGACAGACATCTCTGCACTTGGGGCTTGGAACAGAACTGGCAAG 1200
Db 1081 AGGCCCCCTAGCCAGACAGACATCTCTGCACTTGGGGCTTGGAACAGAACTGGCAAG 1140
QY 1201 AGGCAAGAGTCACTGAGGCGCTCTGTACCCAGGACCTTCCCTGCTGCTCCCTCTCC 1260
Db 1141 AGGCAAGAGTCACTGAGGCGCTCTGTACCCAGGACCTTCCCTGCTGCTCCCTCTCC 1200
QY 1261 CGCCAGACTGTAGAAAATGACACTGTGCCCCAGCCCGAGCTTGGCAGCCCGAGCCGGG 1320
Db 1201 CGCCAGACTGTAGAAAATGACACTGTGCCCCAGCCCGAGCTTGGCAGCCCGAGCCGGG 1260
QY 1321 GTGAGCATGGGCTGGCCACCTCTCTCTTTGTGAGGCTTCCAGCTTCAGGAGGCCA 1380
Db 1261 GTGAGCATGGGCTGGCCACCTCTCTCTTTGTGAGGCTTCCAGCTTCAGGAGGCCA 1320
QY 1381 AGGCTTCTCTCCCAACCGCCCTCCCAAGGCGCTCGGAGCTCAGGTGGCCCGAGT 1440
Db 1321 AGGCTTCTCTCCCAACCGCCCTCCCAAGGCGCTCGGAGCTCAGGTGGCCCGAGT 1380
QY 1441 TCAATCTCCGCTGCTGTGCGCCCTTACCTTCCAGCGCTCCAGTCTCTGGCAGT 1500
Db 1381 TCAATCTCCGCTGCTGTGCGCCCTTACCTTCCAGCGCTCCAGTCTCTGGCAGT 1440
QY 1501 TCTGGAATGAAGGGTTCTGGTGGCCCAACCTGCTCAAGGGCGAGGTCAGGAGTGGGG 1560
Db 1441 TCTGGAATGAAGGGTTCTGGTGGCCCAACCTGCTCAAGGGCGAGGTCAGGAGTGGGG 1500
QY 1561 GCGCTGAGTAGGACTCAGGCGCATGCTGCCCCCTCATCTTCAATCAACCCCACTT 1620
Db 1501 GCGCTGAGTAGGACTCAGGCGCATGCTGCCCCCTCATCTTCAATCAACCCCACTT 1560
QY 1621 AGTTTCTCAGAGCAATCTTCTAGTCTCAAGGGCTAGCATCCTCAGGAGCCAGGCCG 1680
Db 1561 AGTTTCTCAGAGCAATCTTCTAGTCTCAAGGGCTAGCATCCTCAGGAGCCAGGCCG 1620
QY 1681 GCGCAATCCCTCTCTGTCAGAGCTGTCACTTCCGCTGCTGCTGCTGCTGCTG 1740
Db 1621 GCGCAATCCCTCTCTGTCAGAGCTGTCACTTCCGCTGCTGCTGCTGCTGCTG 1680
QY 1741 GTGAGCAGAGTGGAGCTGGGGCGCTGGAGAGCCCGCGCCCTGCACTTCCCTGACC 1800
Db 1681 GTGAGCAGAGTGGAGCTGGGGCGCTGGAGAGCCCGCGCCCTGCACTTCCCTGACC 1740
QY 1801 CGTCTAATATAAATATAGAGATGTGTCTATGCTG 1837
Db 1741 CGTCTAATATAAATATAGAGATGTGTCTATGCTG 1777

RESULT 10
AAI59696/c
ID AAI59696 standard; cDNA; 1868 BP.
XX AC AAI59696;
XX AC
XX AC
DT 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 3685.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.

XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.
XX PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSB-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AAM40540.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
PS Claim 1; SEQ ID NO 3685; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Actinin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX SQ Sequence 1868 BP; 347 A; 536 C; 607 G; 378 T; 0 U; 0 Other;

Query Match 92.2%; Score 1694; DB 4; Length 1868;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1784; Conservative 0; Mismatches 0; Indels 60; Gaps 3;
QY 3 GAGTGGAGATGGCGCGCGGCTCAGGGGGCGGGGGAGCCCGTAGAAC 62
Db 1805 GAGTGGAGATGGCGCGCGGCTCAGGGGGCGGGGGAGCCCGTAGAAC 1746
QY 63 GAGGGGGTCCGCGCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTCCGAC 122
Db 1745 GAGGGGGTCCGCGCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTCCGAC 1686
QY 123 GTGGGGCCCGCTACACGACAGTTCAGTACATCGCGGAGGGCGGTACGGCATGTCAGC 182
Db 1685 GTGGGGCCCGCTACACGACAGTTCAGTACATCGCGGAGGGCGGTACGGCATGTCAGC 1626
QY 183 TCGGCTTATACCACTGCGCGGAGCTCGGCTGCGCATCAAGAGATCAGCCCTTCGAA 242
Db 1625 TCGGCTTATACCACTGCGCGGAGCTCGGCTGCGCATCAAGAGATCAGCCCTTCGAA 1566
QY 243 CATCAGACCTTACTGCCAGCCAGCTCCCGGAGATCCAGATCTCTGCTGCGCTTCGCGCAT 302
Db 1565 CATCAGACCTTACTGCCAGCCAGCTCCCGGAGATCCAGATCTCTGCTGCGCTTCGCGCAT 1506

QY 303 GAGATGTCATCGGCATCCGAGACATTTCTGCGGGCTCCACCTGGAAGCCATGAGAGAT 362
Db |||||
1505 GAGAAATGTCATCGGCATCCGAGACATTTCTGCGGGCTCCACCTGGAAGCCATGAGAGAT 1446
QY 363 GTCTACATTTGTGAGGACCTGATGAGACTGACCTGTACAAAGTTGCTGAAAAGCCAGCAG 422
Db |||||
1445 GTCTACATTTGTGAGGACCTGATGAGAGCTGACCTGTACAAAGTTGCTGAAAAGCCAGCAG 1386
QY 423 CTGAGCAATGACCATATCTGCTACTTTCTCTACAGATCTGCGGGCCCTCAAGTACATC 482
Db |||||
1385 CTGAGCAATGACCATATCTGCTACTTTCTCTACAGATCTGCGGGCCCTCAAGTACATC 1326
QY 483 CACTCCGCGCAAGCTGTCCACCGAGATCTAAAGCCCTCAACCTGTCTCATCAACACACC 542
Db |||||
1325 CACTCCGCGCAAGCTGTCCACCGAGATCTAAAGCCCTCAACCTGTCTCATCAACACACC 1266
QY 543 TGCAGCTTTAAGATTTGTGATTTTGGCTTGGCCCGGATTCGCGATCCTGAGCATGACACC 602
Db |||||
1265 TGCAGCTTTAAGATTTGTGATTTTGGCTTGGCCCGGATTCGCGATCCTGAGCATGACACC 1206
QY 603 ACCGGCTTCTGACGAGTATGTGGCTACGCGCTGTGTACGGGCCCCAGAGATCATGCTG 562
Db |||||
1205 ACCGGCTTCTGACGAGTATGTGGCTACGCGCTGTGTACGGGCCCCAGAGATCATGCTG 1146
QY 663 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTGGCTGAG 722
Db |||||
1145 AACTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTGGCTGAG 1086
QY 723 ATGCTCTCTAAACGGGCCATCTTTCCCTGGCAAGCACTACTCTGGATCAGCTCAACCAATT 782
Db |||||
1085 ATGCTCTCTAAACGGGCCATCTTTCCCTGGCAAGCACTACTCTGGATCAGCTCAACCAATT 1026
QY 783 CTGGGATCTCTGGGCTCCCATCCAGGAGGACCTGAATGTATCATCAACATGAAGGCC 842
Db |||||
1025 CTGGGATCTCTGGGCTCCCATCCAGGAGGACCTGAATGTATCATCAACATGAAGGCC 966
QY 843 CGAAACTTACTACAGTCTCTGCCCTCCAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC 902
Db |||||
965 CGAAACTTACTACAGTCTCTGCCCTCCAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCC 906
QY 903 AAGTCAGACTCCAAAGCCCTTGACCTGGACCGGATGTTAACTTTAACCCCAATAAA 962
Db |||||
905 AAGTCAGACTCCAAAGCCCTTGACCTGGACCGGATGTTAACTTTAACCCCAATAAA 846
QY 963 CGGATCAGTGTGAGAGAGGCTGGCTCACCCTACTCTGAGCAGTACTATGACCCGAGC 1022
Db |||||
845 CGGATCAGTGTGAGAGAGGCTGGCTCACCCTACTCTGAGCAGTACTATGACCCGAGC 786
QY 1023 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTTACCTAAG 1082
Db |||||
785 GATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTTACCTAAG 726
QY 1083 GAGCGGCTGAAGGAGCTCATTTTCCAGGAGACAGCAGCTTCCAGCCCGAGTGTGGAG 1142
Db |||||
725 GAGCGGCTGAAGGAGCTCATTTTCCAGGAGACAGCAGCTTCCAGCCCGAGTGTGGAG 566
QY 1143 GCCCCCTAGCCAGACAGACATCTCTGCACCTTGGGGCTTGGNAACAGAACTGCAAGAG 1202
Db |||||
665 GCCCCCTAGCCAGACAGACATCTCTGCACCTTGGGGCTT----- 626
QY 1203 GCAAGAGGTCACTGAGGGCTCTGTTCACCCAGGACCTGCTCTGCTGCTGCCCTCTCCCG 1262
Db |||||
625 -----GACCTGCTCTGCTGCTGCCCTCTCCCG 597
QY 1263 CCAGACTGTTAGAAAATGGACACTGTGCCAGCCCGACCTTGGCAGCCCGAGCCGGGT 1322
Db |||||
596 CCAGACTGTTAGAAAATGGACACTGTGCCAGCCCGACCTTGGCAGCCCGAGCCGGGT 537
QY 1323 GGAGCATGGGCTGGCCACTCTCTCTTGTGTGAGGCTTCCAGCTTCAGGC-----AG 1376
Db |||||
536 GGAGCATGGGCTGGCCACTCTCTCTTGTGTGAGGCTTCCAGCTTCCAGGCGGCCAAG 477
QY 1377 GCCAAGGCTTCTCTCTCCCGCCAGCCGCTCTCCCGAGGCTCGGGGCTCGGAGCTCAGGTGGCC 1436

Db 476 GCCAAGGCTTCTCTCTCCCGCCCTCCCGACGGGGCTCGGGAGCTCAGGTGGCC 417
QY 1437 CAGTTCAATCTCCC---GCTGCTGCTGCTGGCCCTTACCTTCCCAGGGTCCGAGTCTC 1493
Db 416 CAGTTCAATCTCCCCCTGCTGCTGCTGGCCCTTACCTTCCCAGGGTCCGAGTCTC 357
QY 1494 TGGCAGTTCTGGAATGGAAGGGTTCTGCTGCTGCCCAACCTGCTGAAGGGCAGAGTGGAG 1553
Db 356 TGGCAGTTCTGGAATGGAAGGGTTCTGCTGCTGCCCAACCTGCTGAAGGGCAGAGTGGAG 297
QY 1554 GGTGGGGGGCGCTGAGTAGGGACTCAGGGCCATGCTGCCGCCCTCATCTCATTCATAAACC 1613
Db 296 GGTGGGGGGCGCTGAGTAGGGACTCAGGGCCATGCTGCCGCCCTCATCTCATTCATAAACC 237
QY 1614 CCACCCTAGTTTCCCTGAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGC 1673
Db 236 CCACCCTAGTTTCCCTGAGGAACATTCCTTAGTCTCAAGGCTAGCATCCCTGAGGAGC 177
QY 1674 CAGGCGGGCGGAATCCCTCCTCTCAAAAGCTGTCACCTTCGCGTGCCCTCGCTGCTTCT 1733
Db 176 CAGGCGGGCGGAAATCCCTCCTCTCTCAAGCTGTCACCTTCGCGTGCCCTCGCTGCTTCT 117
QY 1734 GTGCTGTGTGAGCAGAGTGGAGCTGGGGCGGTGGAGAGCCCGGCCCTTGCACCTTC 1793
Db 116 GTGCTGTGTGAGCAGAGTGGAGCTGGGGCGGTGGAGAGCCCGGCCCTTGCACCTTC 57
QY 1794 CTTGACCCGCTTAATATATATATATATAGAGATGTGTCTATGGCTG 1837
Db 56 CTTGACCCGCTTAATATATATATATATAGAGATGTGTCTATGGCTG 13

RESULT 11
AA157910
ID AA157910 standard; cDNA; 1869 BP.
XX
AC AA157910;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 113.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00682191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
(HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX

WPI; 2001-442253/47.
P-PSDB: AAM38754.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Claim 1; SEQ ID NO 113; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA038642-AA042213) with neurotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

Sequence 1869 BP: 378 A; 608 C; 535 G; 348 T; 0 U; 0 Other;

Query Match 91.4%; Score 1679.8; DB 4; Length 1869;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 2; Indels 61; Gaps 4;

Qy	3	GGAGTGGAGATGGCGCGCGCGCGCTCAGGGGGCGCGGGGCGGGGAGCCCCGTAGAACCC	62
Db	64	GGAGTGGAGATGGCGCGCGCGCGCTCAGGGGGCGCGGGGCGGGGAGCCCCGTAGAACCC	123
Qy	63	GAGGGGTCGCGCCCGGGGGTCCC CGGGGAGGTGGAGATGGTGAAGGGGACGCGCTTCGAC	122
Db	124	GAGGGGTCGCGCCCGGGGGTCCC CGGGGAGGTGGAGATGGTGAAGGGGACGCGCTTCGAC	183
Qy	123	GTGGGCCCGCGCTACAACGAGTTGCAGTATACGCGCAGGGCGGTACGGGATGGTTCAGC	182
Db	184	GTGGGCCCGCGCTACAACGAGTTGCAGTATACGCGCAGGGCGGTACGGGATGGTTCAGC	243
Qy	183	TCGGCCTATGACCACTGTCGGCGAAGACTCGCTGGGCCATCAAGAGATCAGGCCCTTCGAA	242
Db	244	TCGGCCTATGACCACTGTCGGCGAAGACTCGCTGGGCCATCAAGAGATCAGGCCCTTCGAA	303
Qy	243	CATCAGACCTACTTGCACGCGCACGCTCCGGGAGATCCAGATCTCTGCTCGCTTCGCGCAT	302
Db	304	CATCAGACCTACTTGCACGCGCACGCTCCGGGAGATCCAGATCTCTGCTCGCTTCGCGCAT	363
Qy	303	GAGAAATGTCATCGGCATCCGAGACATTCCTGGGGGGTCCACCTGGGAAGCCATGAGAGAT	362
Db	364	GAGAAATGTCATCGGCATCCGAGACATTCCTGGGGGGTCCACCTGGGAAGCCATGAGAGAT	423
Qy	363	GTCTACATTTGTGCAGGACCTGATGAGAGATGACCTGTACAAGTTGCTCAAAAGCCAGCAG	422
Db	424	GTCTACATTTGTGCAGGACCTGATGAGAGATGACCTGTACAAGTTGCTCAAAAGCCAGCAG	483
Qy	423	CTGAGCAATGACCATATCTGTACTTCTCTACACAGATCTCTGGGGGGCTCAAGTACATC	482
Db	484	CTGAGCAATGACCATATCTGTACTTCTCTACACAGATCTCTGGGGGGCTCAAGTACATC	543
Qy	483	CACCTCGGCGAAGTGTCTCCACCGAGATCTAAAGCCCTCGAACCTGCTCATCAACACACC	542
Db	544	CACCTCGGCGAAGTGTCTCCACCGAGATCTAAAGCCCTCGAACCTGCTCATCAACACACC	603
Qy	543	TGCGACCTTTAGATTTGTGATTTCCGCTCGGCCCGGATTTGCCGATCTCTGAGCATCACAC	602
Db	604	TGCGACCTTTAGATTTGTGATTTCCGCTCGGCCCGGATTTGCCGATCTCTGAGCATCACAC	663
Qy	603	ACCGGCTTTCCTGACGGAGTATGTGGCTACGCGCTCGGTACCCGGGGCCCCAGAGATCATGCTG	662
Db	664	ACCGGCTTTCCTGACGGAGTATGTGGCTACGCGCTCGGTACCCGGGGCCCCAGAGATCATGCTG	723

Qy	1733	TGTTGTGGTGGAGCAGAACTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTGCCACCT	1792
Db	1753	TGTTGTGGTGGAGCAGAACTGGAGCTGGGGGGCTGGAGAGCCCGCGCCCTGCCACCT	1812
Qy	1793	CCCTGACCCGCTCTAATATATAATAATATAGAGATGTGTCTATGGCTG	1837
Db	1813	CCCTGACCCGCTCTAATATATAATAATATAGAGATGTGTCTATGGCTG	1857
RESULT 12			
ID	ADW12902		
XX	ADW12902 standard; cDNA; 1726 BP.		
AC	ADW12902;		
DT	07-APR-2005 (first entry)		
XX	Human SNAPK3V2 variant gene.		
DE	DNA purification; diagnosis; cancer; mitogen-activated protein kinase;		
XX	gene; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	12..1091	
FT	/tag= a	/product= "SNAPK3V2 protein"	
XX	US2005013817-A1.		
PN	20-JAN-2005.		
XX			
PD	18-JUL-2003; 2003US-00623108.		
XX			
PF	18-JUL-2003; 2003US-00623108.		
XX			
PR	(DAIK/) DAI K.		
XX			
PA	Dai K;		
XX			
PI	WPI; 2005-080923/09.		
DR	P-PSDB; ADW12903.		
XX			
PT	New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing		
PT	diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene		
PT	in a mammal.		
XX			
PS	Claim 4; SEQ ID NO 3; 55pp; English.		
XX			
CC	The invention relates to novel isolated polypeptide (I) comprising an		
CC	amino acid sequence selected from sequences comprising 335 or 359 amino		
CC	acids (ADW12901 or ADW12903) or its fragments. The polypeptides and		
CC	polynucleotides are useful for diagnosing diseases, e.g. cancers,		
CC	associated with the deficiency of the SNAPK3 gene in a mammal. The		
CC	fragments of the polypeptides and polynucleotides can also be used as		
CC	primers or probes. This sequence corresponds to the cDNA encoding the		
CC	SNAPK3V1 variant protein.		
XX			
SQ	Sequence 1726 BP; 339 A; 562 C; 491 G; 334 T; 0 U; 0 Other;		
Query Match 86.8%; Score 1595; DB 14; Length 1726;			
Best Local Similarity 94.0%; Pred. No. 0;			
Matches 1726; Conservative 0; Mismatches 0; Indels 111; Gaps 2;			
Qy	1	GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGGCGGGGCGGGGAGCCCCGTAGAA	60
Db	1	GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGGCGGGGCGGGGAGCCCCGTAGAA	60
Qy	61	CCGAGGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTTCG	120
Db	61	CCGAGGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTTCG	120

QY	121	ACGTGGGCGCGCTACACGAGTTGCACTACATCGCGAGGGCGCGTACGSCATGGTCA	180
Db	121	ACGTGGGCGCGCTACACGAGTTGCACTACATCGCGAGGGCGCGTACGSCATGGTCA	180
QY	181	GCTCGGCTATGACCAACGTTGGCAAGACTCGGTTGGCCATCAAGAAGATCAGCCCTTCG	240
Db	181	GCTCGGCTATGACCAACGTTGGCAAGACTCGGTTGGCCATCAAGAAGATCAGCCCTTCG	240
QY	241	RACATCAGACCTACTGCGAGCGCACGCTCCGGAGATCCAGATCTCTGTGGCTTCCGCC	300
Db	241	RACATCAGACCTACTGCGAGCGCACGCTCCGGAGATCCAGATCTCTGTGGCTTCCGCC	300
QY	301	ATGAGAATGTCATCGGCATCCGAGACATTTCTGCGGGCGTCCACCTCGAAGCCATGAGAG	360
Db	301	ATGAGAATGTCATCGGCATCCGAGACATTTCTGCGGGCGTCCACCTCGAAGCCATGAGAG	360
QY	361	ATGTTCTAATTGTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGC	420
Db	361	ATGTTCTAATTGTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCCAGC	420
QY	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACAGATCTCGGGGCGCTCAAGTACA	480
Db	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACAGATCTCGGGGCGCTCAAGTACA	480
QY	481	TCCACTCCGCGCAACGTTGATTTGCGGCTGGCCCGGATTCGCCATCTGAGCATGACC	540
Db	481	TCCACTCCGCGCAACGTTGATTTGCGGCTGGCCCGGATTCGCCATCTGAGCATGACC	540
QY	541	CCTGCGACCTTAAGATTGTTGATTTGCGGCTGGCCCGGATTCGCCATCTGAGCATGACC	600
Db	541	CCTGCGACCTTAAGATTGTTGATTTGCGGCTGGCCCGGATTCGCCATCTGAGCATGACC	600
QY	601	ACACCGGCTTCTGACGAGTATGTTGCTACCGCTGGTACCGGGCCCGAGAGATCATGC	660
Db	601	ACACCGGCTTCTGACGAGTATGTTGCTACCGCTGGTACCGGGCCCGAGAGATCATGC	660
QY	661	TGAATCTCCAAGGGCTATACCAAGTCCATCGACATCTGTTGTTGGCTGCACTTCTGGCTG	720
Db	661	TGAATCTCCAAGGGCTATACCAAGTCCATCGACATCTGTTGTTGGCTGCACTTCTGGCTG	720
QY	721	AGATGCTCTTAACCGGCGCATCTTCCCTGGCAAGCAGCTACCTGGATCAGCTCAACCA	780
Db	721	AGATGCTCTTAACCGGCGCATCTTCCCTGGCAAGCAGCTACCTGGATCAGCTCAACCA	780
QY	781	TTCTGGGATCTCTGGGCTCCCCATCCAGGAGGACCTGAATTGTATCATCAACCAAGG	840
Db	781	TTCTGGGATCTCTGGGCTCCCCATCCAGGAGGACCTGAATTGTATCATCAACCAAGG	840
QY	841	CCCGAAACTACCTACAGTCTCTGCCCTCCAAAGCAAGTGGCTTGGGCAAGCTTTTCC	900
Db	841	CCCGAAACTACCTACAGTCTCTGCCCTCCAAAGCAAGTGGCTTGGGCAAGCTTTTCC	900
QY	901	CCAAAGTCAAGCTCCAAAGCCCTTGACCTGTGACCGGATGTTAACTTTAACCCCAATA	960
Db	901	CCAAAGTCAAGCTCCAAAGCCCTTGACCTGTGACCGGATGTTAACTTTAACCCCAATA	960
QY	961	RACGATCACAGTGGAGGAGGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA	1020
Db	961	RACGATCACAGTGGAGGAGGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA	1020
QY	1021	CGGATGAGCCAGTGGCGGAGGAGCCCTTACCTTCGCCATGGAGCTGGATGACCTTACCTA	1080
Db	970	-----CAGTGGCCGAGGAGCCCTTACCTTCGCCATGGAGCTGGATGACCTTACCTA	1020
QY	1081	AGGAGCGGCTGAAGAGCTCATCTTCAGGAGACAGCAGCTTCCAGCCCGGAGTGTCTGG	1140
Db	1021	AGGAGCGGCTGAAGAGCTCATCTTCAGGAGACAGCAGCTTCCAGCCCGGAGTGTCTGG	1080
QY	1141	AGGCCCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCGCTGGAACAGACTGGCAAG	1200
Db	1081	AGGCCCCCTAGCCAGACAGACATCTCTGACCCCTGGGGCGCT-----	1122
QY	1201	AGGCAAGAGGTCACTGAGGGGCTCTGTGTACCCAGGAGACCTGTGCTCTGTGCTCTCTCC	1260

Db 661 TGAAGTCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 720
Qy 721 AGATGCTCTTAACCGGCCCATTTCTCCCTGGGAGCACTACCTCGGATCAGCTCAACCA 780
Db 721 AGATGCTCTTAACCGGCCCATTTCTCCCTGGGAGCACTACCTCGGATCAGCTCAACCA 780
Qy 781 TTCTGGGCATCTCTGGGCTCCCATCCAGGAGGACCTGAATTGTATCATCAACATCAAGG 840
Db 781 TTCTG----- 785
Qy 841 CCCGAACTACTPACAGTCTTGCCCTCCAAGCAAGGTGGCTTGGCCAAAGCTTTTCC 900
Db 786 ----- 785
Qy 901 CCAAGTCAGACTCCAAAGCCCTTGACCTGTGACCGGATGTTAACTTTTACCCCAATA 960
Db 786 -----GCCCTTGACCTGTGACCGGATGTTAACTTTTACCCCAATA 928
Qy 961 AACGGATCAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1020
Db 829 AACGGATCAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 888
Qy 1021 CGGATGAGCCAGTGGCGGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTACTPA 1080
Db 889 CGGATGAGCCAGTGGCGGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTACTPA 948
Qy 1081 AGAGCGGCTGAAGGAGCTCATCTTCAGGAGACAGCAGCTTCCAGCGCGGAGTCTGG 1140
Db 949 AGAGCGGCTGAAGGAGCTCATCTTCAGGAGACAGCAGCTTCCAGCGCGGAGTCTGG 1008
Qy 1141 AGGCCCCCTAGCCAGACAGACATCTTGACCCCTGGGGCTTGGAAACAGACTGGCAAG 1200
Db 1009 AGGCCCCCTAGCCAGACAGACATCTTGACCCCTGGGGCT----- 1050
Qy 1201 AGGCAAGAGTCACTGAGGSCCTCTGTACCCAGGACCTGCTCTGCTGCTGCCCTCTCC 1260
Db 1051 -----GGACCTGCTCTGCTGCCCTCTCC 1077
Qy 1261 CGCCAGACTGTTAGAAATGGACACTGTGCCAGCGCGGACCTTGGACGCCAGCGCGG 1320
Db 1078 CGCCAGACTGTTAGAAATGGACACTGTGCCAGCGCGGACCTTGGACGCCAGCGCGG 1137
Qy 1321 GTGAGCATGGGCTGGCCACCTCTCTCTCTGCTGAGGCTCCAGCTTCAGGAGGCCA 1380
Db 1138 GTGAGCATGGGCTGGCCACCTCTCTCTCTGCTGAGGCTCCAGCTTCAGGAGGCCA 1197
Qy 1381 AGGCTTCTCTCCCAACCGCCCTCCCAACCGGCTCGGAGCTCAGGTGGCCCACT 1440
Db 1198 AGGCTTCTCTCCCAACCGCCCTCCCAACCGGCTCGGAGCTCAGGTGGCCCACT 1257
Qy 1441 TCAATCTCCGCTGCTGCTGGGCGCTTACCTTCCAGCGCTCCAGCTCTCTGGCAGT 1500
Db 1258 TCAATCTCCGCTGCTGCTGGGCGCTTACCTTCCAGCGCTCCAGCTCTCTGGCAGT 1317
Qy 1501 TCTGGAATGAAGGGTCTGGCTGCCCACTGCTGAAGGGCAGAGGTGGAGGGTGGG 1560
Db 1318 TCTGGAATGAAGGGTCTGGCTGCCCACTGCTGAAGGGCAGAGGTGGAGGGTGGG 1377
Qy 1561 GCGCTGAGTAGGAGCTCAGGGGCATGCTGCTGCCCTCATCTCATATCAATCAAAACCCACCT 1620
Db 1378 GCGCTGAGTAGGAGCTCAGGGGCATGCTGCTGCCCTCATCTCATATCAAAACCCACCT 1437
Qy 1621 AGTTTCCCTGAAGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGCGG 1680
Db 1438 AGTTTCCCTGAAGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGCGG 1497
Qy 1681 GGCAGATCCCTCTCCCTGTCAAAGCTGTCTCTGCGTGGCTCGCTGCTCTGTGTGTG 1740
Db 1498 GGCAGATCCCTCTCCCTGTCAAAGCTGTCTCTGCGTGGCTCGCTGCTCTGTGTGTG 1557
Qy 1741 GTGACAGAAAGTGGAGCTGGGGGCGTGGAGAGCGCGGCGCCCTGCTGCACTCCCTGACC 1800
Db 1558 GTGACAGAAAGTGGAGCTGGGGGCGTGGAGAGCGCGGCGCCCTGCTGCACTCCCTGACC 1617

Qy 1801 CGTCTAATATATAATATAGAGATGCTCTATGGCTG 1837
Db 1618 CGTCTAATATATAATATAGAGATGCTCTATGGCTG 1654

RESULT 14

ADH48413
ID ADH48413 standard; DNA; 1699 BP.

XX ADH48413;

DT 25-MAR-2004 (first entry)

XX Human KPP encoding DNA SEQ ID NO:71.

DE human; KPP; kinase; phosphatase; enzyme; cytostatic; anorectic;
XX immunosuppressive; KPP-Antagonist; KPP-Agonist; gene therapy;
KW autoimmune disorder; obesity; cancer; ds; gene.
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 25..1032
FT /*tag= a
FT /product= "KPP"

XX WO2004001008-A2.

XX 31-DEC-2003.

XX 19-JUN-2003; 2003WO-US019660.

XX 21-JUN-2002; 2002US-0390652P.

XX 15-JUL-2002; 2002US-0396196P.

XX (INCY-) INCYTE CORP.

XX Marquis JP, Baughn MR, Tran UK, Hafalia AJA, Kable AE,
PI Emerling BM, Elliott VS, Lindquist EA, Richardson TW, Khare R;
PI Swarnakar A, Lee SY, Ramkumar J, Chawla NK, Becha SD, Mason PM;
PI Hawkins PR, Bulloch SA, Jin P, Bhatia U, Burrill JD, Lee S;
PI Blake JU, Ho A, Zheng W;

XX WPI: 2004-082489/08.

XX P-PSDB; ADH48364.

XX New human kinases and phosphatases (KPP) polypeptide, useful for
PT preparing a composition for treating a disease associated with decreased
PT expression or overexpression of functional KPP e.g., cancer.

XX Claim 5; SEQ ID NO 71; 336pp; English.

XX The invention relates to novel isolated human kinases and phosphatases
CC (KPP) polypeptides. A protein of the invention has cytostatic, anorectic,
CC and immunosuppressive activity, and acts as a KPP-Antagonist, or KPP-
CC Agonist. A polynucleotide of the invention may have a use in gene
CC therapy. The polypeptide is useful for preparing a composition for
CC diagnosing or treating a disease or condition associated with decreased
CC expression or overexpression of functional KPP e.g., autoimmune
CC disorders, obesity or cancer. The sequences shown in ADH48392-ADH48440
CC represent KPP proteins of the invention.

XX Sequence 1699 BP; 348 A; 538 C; 494 G; 319 T; 0 U; 0 Other;

Query Match 78.2%; Score 1437; DB 12; Length 1699;

Best Local Similarity 89.8%; Pred. No. 0;

Matches 1654; Conservative 0; Mismatches 0; Indels 187; Gaps 3;

Qy 1 GAGGAGTGAGATGGCGCGCGGCTCAGGGGGCGGGGCGGGAGCCCGTAGAA 60
Db 14 GAGGAGTGAGATGGCGCGCGGCTCAGGGGGCGGGGCGGGAGCCCGTAGAA 73

QY 61 CCAGGGGGTCCGCCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACGCGTTGC 120
DB 74 CCAGGGGGTCCGCCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACGCGTTGC 133
QY 121 ACCTGGGCGCGCTACACGAGTTGACGTACATCCGCGAGGGGCGGTACGGCATGGTCA 180
DB 134 ACCTGGGCGCGCTACACGAGTTGACGTACATCCGCGAGGGGCGGTACGGCATGGTCA 193
QY 181 GCTCGGCTATGATCCAGCTCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 240
DB 194 GCTCGGCTATGATCCAGCTCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 253
QY 241 AACATCAGACTACTGTCAGCGACGCTCCGAGAGTCCAGATCCTGCTCGCTTCGCGC 300
DB 254 AACATCAGACTACTGTCAGCGACGCTCCGAGAGTCCAGATCCTGCTCGCTTCGCGC 313
QY 301 ATGAGATGTCATCGGATCCGAGACATCTCGGGGGTCCACCTCGAAGCCATGAGAG 360
DB 314 ATGAGATGTCATCGGATCCGAGACATCTCGGGGGTCCACCTCGAAGCCATGAGAG 373
QY 361 ATGTCTACATTTGTCAGGACTGATGAGACTGACCTGTACAAAGTTGCTGAAAAGCCAGC 420
DB 374 ATGTCTACATTTGTCAGGACTGATGAGACTGACCTGTACAAAGTTGCTGAAAAGCCAGC 433
QY 421 AGCTGAGCAATGACCATATCTGCTACTCTCTACAGATCTCTGCGGGGCTCAAGTACA 480
DB 434 AGCTGAGCAATGACCATATCTGCTACTCTCTACAGATCTCTGCGGGGCTCAAGTACA 493
QY 481 TCCACTCCGCAAGGTGCTCCACCGAGATCTTAAGCCCTCCAACTGCTCATCAACACCA 540
DB 494 TCCACTCCGCAAGGTGCTCCACCGAGATCTTAAGCCCTCCAACTGCTCATCAACACCA 553
QY 541 CCTGCGACCTTAAGATTTGATTTCCGCGCTGGCCCGGATTCGCGATCTCTGAGCATGACC 600
DB 554 CCTGCGACCTTAAGATTTGATTTCCGCGCTGGCCCGGATTCGCGATCTCTGAGCATGACC 613
QY 601 ACACCGGCTTCTGACGAGATATGCTGCTACGCGCTGGTACCGGGCCCCAGAGATATGC 660
DB 614 ACACCGGCTTCTGACGAGATATGCTGCTACGCGCTGGTACCGGGCCCCAGAGATATGC 673
QY 661 TGAATCTCAAGGCTATACCAAGTCCATCCGATCTGCTGCTGGGCTGATCTGGCTG 720
DB 674 TGAATCTCAAGGCTATACCAAGTCCATCCGATCTGCTGCTGGGCTGATCTGGCTG 733
QY 721 AGATGCTCTCTAACCGGCCCATCTTCCCTGCGAGCACTACTCTGAGTCAACCCACA 780
DB 734 AGATGCTCTCTAACCGGCCCATCTTCCCTGCGAGCACTACTCTGAGTCAACCCACA 793
QY 781 TTCTGGGCATCTCGGGCTCCCCATCCAGGAGGACCTGAATTTGATCATCAACATGAAG 840
DB 794 TTCTG----- 798
QY 841 CCGAAGACTACCTACAGTCTCTGCCCTCCAAAGCAAGGTGGCTTGGGCCAAGCTTTTCC 900
DB 799 ----- 798
QY 901 CCAAGTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAAACCCCAATA 960
DB 799 -----GCCCTTGACCTTGCTGGACCGGATGTTAACTTTAAACCCCAATA 841
QY 961 AACGGATCACAGTGGAGGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA 1020
DB 842 AACGGATCACAGTGGAGGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA 901
QY 1021 CGGATGAGCCAGTGGCGAGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTACTTA 1080
DB 902 CGGATGAGCCAGTGGCGAGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTACTTA 961
QY 1081 AGGAGCGGCTCAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGG 1140
DB 962 AGGAGCGGCTCAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGTGG 1021
QY 1141 AGGCCCCCTAGCCCCAGACAGACATCTCTGCAACCTCGGGGCTGGAAACAGAACTGGCAAG 1200

DB 1022 AGGCCCCCTAGCCCCAGACAGACATCTCTGACCCCTGGGGCT----- 1063
QY 1201 AGGCAAGAGGTCACTGAGGGCTCTGTCAACCCAGAGACCTGCTCTCTGCTGCCCTCTCC 1260
DB 1064 -----GGACCTGCTCTCTGCTGCCCTCTCC 1090
QY 1261 CGCAGACTGTTAGAAAATGGACACTGTGCGCAGCCCGGACCTTGGCAGCCAGGCGGG 1320
DB 1091 CGCAGACTGTTAGAAAATGGACACTGTGCGCAGCCCGGACCTTGGCAGCCAGGCGGG 1150
QY 1321 GTGGAGCATGGGCTCGGCCACCTCTCTCTTCTGAGGCTCCAGCTTCCAGGAGGCCA 1380
DB 1151 GTGGAGCATGGGCTCGGCCACCTCTCTCTTCTGAGGCTCCAGCTTCCAGGAGGCCA 1210
QY 1381 AGSCCTTCTCCTCCCAACCGGCTCCCAACCGGGGCTCGGGAGCTCAGGTGCCCCAGT 1440
DB 1211 AGSCCTTCTCCTCCCAACCGGCTCCCAACCGGGGCTCGGGAGCTCAGGTGCCCCAGT 1270
QY 1441 TCAATCTCCGCT-----GCTGCTGCTGGGCTCTTACCTTCCCGAGGTTCCAGTCTCTGG 1496
DB 1271 TCAATCTCCGCTGCTGGCTGCTGCTGCGGCTTACCTTCCCGAGGTTCCAGTCTCTGG 1330
QY 1497 CAGTCTCGAATCGAAGGGTCTGCTGCTGCGGCTTCCCAACCTGCTGAAGGCGAGAGTGGAGGT 1556
DB 1331 CAGTCTCGAATCGAAGGGTCTGCTGCTGCGGCTTCCCAACCTGCTGAAGGCGAGAGTGGAGGT 1390
QY 1557 GGGGGCGCTGAGTAGGGACTCAGGGGCCATGCTGCTGCCCTCATCTCATTTCAAACCCCA 1616
DB 1391 GGGGGCGCTGAGTAGGGACTCAGGGGCCATGCTGCTGCCCTCATCTCATTTCAAACCCCA 1450
QY 1617 CCTAGTCTTCCCTGAAGGAACATCTTAGTCTCAAGGGCTAGATCCCTGAGGAGCCAG 1676
DB 1451 CCTAGTCTTCCCTGAAGGAACATCTTAGTCTCAAGGGCTAGATCCCTGAGGAGCCAG 1510
QY 1677 GCGGGCCGAATCCCTCCCTGCTCAAGGCTGCTCACTTCCGGTGGCTGCTGCTGCTG 1736
DB 1511 GCGGGCCGAATCCCTCCCTGCTCAAGGCTGCTCACTTCCGGTGGCTGCTGCTGCTG 1570
QY 1737 TGTGCTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTGCTGCCCTCCCT 1796
DB 1571 TGTGCTGAGCAGAAAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGCCCTGCTGCCCTCCCT 1630
QY 1797 GACCCGCTTAATATATAATATATAGAGATGTGTCTATGGCTG 1837
DB 1631 GACCCGCTTAATATATAATATATAGAGATGTGTCTATGGCTG 1671

RESULT 15

ADH59631
ID ADH59631 standard; DNA; 1140 BP.

XX AC ADH59631;

XX DX 25-MAR-2004 (first entry)

XX DE Erk1 encoding sequence.

XX KW cardiac disease; Raf-1; MEK1; Cardiant; telethonin; &agr-actinin;
XX KW hypertrophy; ds; MEK2; Erk1; Erk2.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX CDS 1..1140

XX FT /*tag= a
XX FT /product= "Erk1"

XX WO2003025205-A2.

XX PN 27-MAR-2003.

XX PD 18-SEP-2002; 2002WO-BP010489.

XX PF

XX 19-SEP-2001; 2001US-0323566P.
PR 24-SEP-2001; 2001US-0324625P.
XX (MEDI-) MEDIGENE AG.
XX Nave B, Roenicke V, Leclair S, Funk M, Reuner B, Brinkmann K;
PI Henkel T;
XX WPI; 2003-371821/35.
DR P-PSDB; ADH59632.
XX Identifying and/or obtaining a compound useful for preventing or treating
PT cardiac diseases, particularly congestive heart failure, comprises
PT quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
PT on the compound.
XX
PS Claim 33; SEQ ID NO 7; 41pp; English.
XX The present invention relates to identifying and obtaining a compound
CC useful in the prevention or treatment of cardiac diseases, comprising
CC quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
CC on the compound. The method is useful for identifying and/or obtaining
CC compounds that may be used in the prevention or treatment of cardiac
CC diseases, particularly congestive heart failure. The compound or protein
CC is also used for the preparation of a pharmaceutical composition for
CC prevention or treatment of a disease related to hypertrophy or impaired
CC or increased activation of telethonin (T-Cap), α actinin, MHC, actin,
CC titin, myomesin, nebulin, tropomyosin, troponin, Erk1/2 and/or MLCK.
CC The protein or antibody which specifically recognizes the
CC activated/phosphorylated form of the above polypeptide, is used for the
CC preparation of a composition for diagnosing a disease or a predisposition
CC for a disease related to hypertrophy or related to impaired or increased
CC activation of Raf-1, MEK1/2 and/or Erk1/2. The present sequence
CC represents Erk1 encoding sequence.
XX
SQ Sequence 1140 BP; 247 A; 357 C; 324 G; 212 T; 0 U; 0 Other;

Query Match 62.0%; Score 1138.4; DB 10; Length 1140;
Best Local Similarity 99.9%; Pred. No. 2.2e-247;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 ATGGCGGCGGCGGCGCTCAGGGGGGCGGGGGGCGGAGCCCGTAGAACCCGAGGGGGTC 71
DB 1 ATGGCGGCGGCGGCGCTCAGGGGGGCGGGGGGCGGAGCCCGTAGAACCCGAGGGGGTC 60
QY 72 GGCGCGGGGGTCCCGGGGGAGGTGGAGATGGTGAAGGGGCGCGGCTTCGACCGTGGGCCG 131
DB 61 GGCGCGGGGGTCCCGGGGGAGGTGGAGATGGTGAAGGGGCGCGGCTTCGACCGTGGGCCG 120
QY 132 CGGTACACGAGTTGCAGTACATCGCGGGAGGGCGGTACGGCATGGTTCAGCTCGGCTAT 191
DB 121 CGGTACACGAGTTGCAGTACATCGCGGGAGGGCGGTACGGCATGGTTCAGCTCGGCTAT 180
QY 192 GACCAGTGGCAAGACTCGGTGGCCATCAAGAAGATCAGCCCTTCGAAATCAGACC 251
DB 181 GACCAGTGGCAAGACTCGGTGGCCATCAAGAAGATCAGCCCTTCGAAATCAGACC 240
QY 252 TACTGCCAGCGCAGCTCCGGGAGATCCAGATCTCTGCTGGCTTCGCCCATGAGATGTC 311
DB 241 TACTGCCAGCGCAGCTCCGGGAGATCCAGATCTCTGCTGGCTTCGCCCATGAGATGTC 300
QY 312 ATCGGATCCGAGACATTCCTGGGGGCTCCACCCCTGGAAGCCATGAGATGTCTACATT 371
DB 301 ATCGGATCCGAGACATTCCTGGGGGCTCCACCCCTGGAAGCCATGAGATGTCTACATT 360
QY 372 GTGCGAGCTGATGAGACTGACCTGTGTAAGTGTCTGAAAGCCAGCGCTGAGCAAT 431
DB 361 GTGCGAGCTGATGAGACTGACCTGTGTAAGTGTCTGAAAGCCAGCGCTGAGCAAT 420
QY 432 GACCATATCTGCTACTTCCTTACAGATCTCTGGGGGCTCAAGTACATCCACTCCGCC 491
DB 421 GACCATATCTGCTACTTCCTTACAGATCTCTGGGGGCTCAAGTACATCCACTCCGCC 480

QY 492 AACGTGCTCCACCGAGATCTAAAGCCCTCCAACCTGCTCATCAACACCACTTGGACCTT 551
DB 481 AACGTGCTCCACCGAGATCTAAAGCCCTCCAACCTGCTCAGCAGACCACTTGGACCTT 540
QY 552 AAGATTTGTGATTTTCGGCTCGCGGATTCGGCATCTCGAGATGACACACCGGCTTC 611
DB 541 AAGATTTGTGATTTTCGGCTCGCGGATTCGGCATCTCGAGATGACACACCGGCTTC 600
QY 612 CTGACGGAGTATGTGGCTACCGGCTGGTACCGGGCCCGAGAGATCATGCTGAATCCCAAG 671
DB 601 CTGACGGAGTATGTGGCTACCGGCTGGTACCGGGCCCGAGAGATCATGCTGAATCCCAAG 660
QY 672 GGTATACCAAGTCCCATCGACATCTGGTCTGGGCTGCTTCTGGCTGAGATGCTCTCT 731
DB 661 GGTATACCAAGTCCCATCGACATCTGGTCTGGGCTGCTTCTGGCTGAGATGCTCTCT 720
QY 732 AACCGGCGCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCACTTCTGGGCTTC 791
DB 721 AACCGGCGCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCACTTCTGGGCTTC 780
QY 792 CTGGGCTCCCATCCAGGAGACCTGATTTGTATCATCAACATGAAGCCCGAACTAC 851
DB 781 CTGGGCTCCCATCCAGGAGACCTGATTTGTATCATCAACATGAAGCCCGAACTAC 840
QY 852 CTACAGTCTCTGCCCTCCAAAGACCAAGGTGGCTTGGGCGCAAGCTTTTCCCAAGTCAGAC 911
DB 841 CTACAGTCTCTGCCCTCCAAAGACCAAGGTGGCTTGGGCGCAAGCTTTTCCCAAGTCAGAC 900
QY 912 TCCAAAGCCCTTGACCTCTGGACCGGATGTTAACTTTAACCCCAATAAAGCGATCACA 971
DB 901 TCCAAAGCCCTTGACCTCTGGACCGGATGTTAACTTTAACCCCAATAAAGCGATCACA 960
QY 972 GTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCGGAGTGAAGCA 1031
DB 961 GTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCGGAGTGAAGCA 1020
QY 1032 GTGGCGGAGGAGCCCTTCACTTCCCATGGAGCTGGATGACCTTACCTAAGGAGGGCTG 1091
DB 1021 GTGGCGGAGGAGCCCTTCACTTCCCATGGAGCTGGATGACCTTACCTAAGGAGGGCTG 1080
QY 1092 AAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTCTGGAGGCCCTTAG 1151
DB 1081 AAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTCTGGAGGCCCTTAG 1140

Search completed: February 7, 2006, 12:49:02
Job time : 1133.96 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:29:45 ; Search time 7228.32 Seconds
(without alignments)
11890.431 Million cell updates/sec

Title: US-10-623-108-5
Perfect score: 1837
Sequence: 1 gagagtgagatggcggcg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1712.4	93.2	1775	4	CR620058 full-leng
2	1699.4	92.5	1762	4	CR621455 full-leng
3	1694.4	92.2	1764	4	CR606522 full-leng
4	1688.4	91.9	1751	4	CR595737 full-leng
5	1681.4	91.5	1751	4	CR597147 full-leng
6	1673.4	91.1	1854	4	CR603463 full-leng
7	1666.4	90.7	1729	4	CR621977 full-leng
8	1644.4	89.5	1707	4	CR621688 full-leng
9	1615.4	87.9	1678	4	CR600245 full-leng
10	1577.4	85.9	1718	4	CR596064 full-leng
11	1549.4	84.3	1612	4	CR607536 full-leng
12	1034.2	56.3	1130	1	AL530655 AL530655
13	1008	54.9	1201	1	AL530654 AL530654
14	960.2	52.3	1016	1	AL530980 AL530980
15	945.2	51.5	1108	5	EX445445 EX445445
16	934.8	50.9	1013	1	AL515713 AL515713
17	933.8	50.8	946	1	AL526297 AL526297
18	913.8	49.7	989	1	AL583198 AL583198
19	908.6	49.5	985	5	BX398410 BX398410
20	906	49.3	949	1	AL549418 AL549418
21	898.4	48.9	986	1	AL529832 AL529832
22	883.8	48.1	1089	3	BM548079 BM548079

23	882.2	48.0	910	5	BX386989
24	876.8	47.7	898	1	AL559377
25	875	47.6	875	10	AY403905
26	872.6	47.5	1131	3	BM462994
27	867.2	47.2	902	1	AL533916
28	867	47.2	1083	5	BX333599
29	862.6	47.0	921	7	CR999816
30	859	46.8	1036	3	BQ66864
31	856.2	46.6	909	7	CR981539
32	855.2	46.6	882	5	BQ890149
33	854.6	46.5	1015	2	BI225074
34	853.2	46.4	963	1	AL526262
35	850.4	46.3	968	5	BX354177
36	845	46.0	877	1	AL558763
37	843	45.9	843	1	AL516647
38	842.8	45.9	847	10	AY403906
39	837.4	45.6	1053	3	BM468615
40	832.8	45.3	864	1	AL576621
41	830.4	45.2	1067	1	AL544810
42	829.6	45.2	971	3	BI768741
43	825.6	44.9	1128	1	AL583197
44	820.8	44.7	1080	5	EX445446
45	819.4	44.6	847	3	BI768732

ALIGNMENTS

RESULT 1
CR620058

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

Features

Source

Origin

Query Match

Best Local Similarity

Matches 1774; Conservative

Score 93.2%;

Pred. No. 0;

Mismatches 0;

Indels 1;

Gaps 1;

Length 1775;

DB 4;

Score 1712.4;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

CR620058 1775 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DA001YE10 of Neuroblastoma of Homo sapiens (human).

Accession CR620058.1 GI:50500865

Version H7C; CNSLT_cDNA.

Keywords Homo sapiens (human)

Source Homo sapiens

Organism Homo sapiens

Reference 1 (bases 1 to 1775)

Authors Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Title Full-length cDNA libraries and normalization

Journal Unpublished

Comment Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Reference 2 (bases 1 to 1775)

Authors Genoscope.

Title Direct Submission

Journal Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

Comment 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

Features Location/Qualifiers

1..1775

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DA001YE10"

/tissue_type="Neuroblastoma"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 93.2%;

Best Local Similarity 97.2%;

Matches 1774; Conservative

Score 1712.4;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

Score 97.2%;

QY 1 GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAA 60
DB 1 GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAA 60
QY 61 CCGAGGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTCC 120
DB 61 CCGAGGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTCC 120
QY 121 ACCTGGGCGCGCTACACGAGTTGAGTACATCGGCGAGGGCGGTAGGGATGGTCA 180
DB 121 ACCTGGGCGCGCTACACGAGTTGAGTACATCGGCGAGGGCGGTAGGGATGGTCA 180
QY 181 GCTCGGCTATGACCAAGCTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCGCTTCG 240
DB 181 GCTCGGCTATGACCAAGCTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCGCTTCG 240
QY 241 AACATCAGACCTACTGCGCAGCGCACGCTCCGGGAGATCCAGATCTCTGCTCCGCG 300
DB 241 AACATCAGACCTACTGCGCAGCGCACGCTCCGGGAGATCCAGATCTCTGCTCCGCG 300
QY 301 ATGAGAAATGTCATCGGATCCGAGACATTTCTGCGGGGTCACCTTGGGAAGCCATGAGAG 360
DB 301 ATGAGAAATGTCATCGGATCCGAGACATTTCTGCGGGGTCACCTTGGGAAGCCATGAGAG 360
QY 361 ATGCTACATTTGTCAGGACCTCATGAGACTGACCTGTACAGTTGCTGAAAGCCAGC 420
DB 361 ATGCTACATTTGTCAGGACCTCATGAGACTGACCTGTACAGTTGCTGAAAGCCAGC 420
QY 421 AGCTGAGCAATGACCAATATCTGCTACTTCTTACCAGATCTCTGCGGGGCTCAAGTACA 480
DB 421 AGCTGAGCAATGACCAATATCTGCTACTTCTTACCAGATCTCTGCGGGGCTCAAGTACA 480
QY 481 TCCACTCCGCGCAACGCTGCTCCACCGAGATCAAAGCCCTCCAACTGCTCATCAACACCA 540
DB 481 TCCACTCCGCGCAACGCTGCTCCACCGAGATCAAAGCCCTCCAACTGCTCATCAACACCA 540
QY 541 CTTGCGACCTTAAGATTGTTGATTTTCGGCTGCGCCGAGTTGCGGATCTTGAGCATGACC 600
DB 541 CTTGCGACCTTAAGATTGTTGATTTTCGGCTGCGCCGAGTTGCGGATCTTGAGCATGACC 600
QY 601 ACACCGGCTTCCTGACGGAGTATGGCTACCGGCTGGTACCGGGGCCCGAGAGATCATGC 660
DB 601 ACACCGGCTTCCTGACGGAGTATGGCTACCGGCTGGTACCGGGGCCCGAGAGATCATGC 660
QY 661 TGAATCTCAAGGGGTATACCAAGTTCATCGACATCTGGTCTGTTGGGCTGATCTGGCTG 720
DB 661 TGAATCTCAAGGGGTATACCAAGTTCATCGACATCTGGTCTGTTGGGCTGATCTGGCTG 720
QY 721 AGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGTCAACACCA 780
DB 721 AGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGTCAACACCA 780
QY 781 TTCTGGGATCTCTGGGCTCCCATCCAGGAGGACCTGAATTTGATCATCAATGAAGG 840
DB 781 TTCTGGGATCTCTGGGCTCCCATCCAGGAGGACCTGAATTTGATCATCAATGAAGG 840
QY 841 CCCGAAATACCTACAGTCTCTGCGCTTCAAGACCAAGTGGCTTGGGCCAAGCTTTTCC 900
DB 841 CCCGAAATACCTACAGTCTCTGCGCTTCAAGACCAAGTGGCTTGGGCCAAGCTTTTCC 900
QY 901 CCAAGTCAGACTCCAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAACCCCAATA 960
DB 901 CCAAGTCAGACTCCAAGCCCTTGACCTGCTGGACCGGATGTTAACTTTAACCCCAATA 960
QY 961 AACGGATCAAGTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1020
DB 961 AACGGATCAAGTGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1020
QY 1021 CGGATGAGCCAGTGGCGGAGGAGCCCTTACCTTGGCCATGGAGCTGGATGACCTACCTA 1080
DB 1021 CGGATGAGCCAGTGGCGGAGGAGCCCTTACCTTGGCCATGGAGCTGGATGACCTACCTA 1080
QY 1081 AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCACGCTTCCAGCGCGGAGTGTGG 1140

DB 1081 AGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCACGCTTCCAGCGCGGAGTGTGG 1140
QY 1141 AGGCGCCCTTAGCCACAGACATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAG 1200
DB 1141 AGGCGCCCTTAGCCACAGACATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAG 1182
QY 1201 AGGCAAGAGGTCACTGAGGGCTCTGTCAACCAGGACCTTGCCCTCTGCTGCTGCCCTCTCC 1260
DB 1183 -----GGACTGCTCTCTGCTGCCCTCTCTCC 1209
QY 1261 CGCCAGACTGTTAGAAAATGGAACACTGTGCCAGCCCGAACCTTGGCAGCCAGGCCCGGG 1320
DB 1210 CGCCAGACTGTTAGAAAATGGAACACTGTGCCAGCCCGAACCTTGGCAGCCAGGCCCGGG 1269
QY 1321 GTGGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTCCAGCTTCAAGCAGGCCA 1380
DB 1270 GTGGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTCCAGCTTCAAGCAGGCCA 1329
QY 1381 AGGCTTCTCTCCACCCCGCTTACCTTCCCAGCGCTCGGGAGCTCAGGTGGCCCCAGT 1440
DB 1330 AGGCTTCTCTCTCCACCCCGCTTACCTTCCCAGCGCTCGGGAGCTCAGGTGGCCCCAGT 1389
QY 1441 TCAATCTCCCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1390 TCAATCTCCCGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449
QY 1501 TCTGGAATGGAAGGGTTCTGGCTGGCCCAACCTGTGTAAGGCGAGGTGGAGGTGGGG 1560
DB 1450 TCTGGAATGGAAGGGTTCTGGCTGGCCCAACCTGTGTAAGGCGAGGTGGAGGTGGGG 1509
QY 1561 GCGCTGTAGTAGGACTCAGGGCCATGCTGCTGCCCTCATCTCATTTCAAAACCCCAACCT 1620
DB 1510 GCGCTGTAGTAGGACTCAGGGCCATGCTGCTGCCCTCATCTCATTTCAAAACCCCAACCT 1569
QY 1621 AGTTTCCCTGGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCTCTGAGGAGCCAGGCCG 1680
DB 1570 AGTTTCCCTGGAAGGAACATTCCTTAGTCTCAAGGCTAGCATCTCTGAGGAGCCAGGCCG 1629
QY 1681 GCGGAAATCCCTCTCTGTCAAGCTGTCACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
DB 1630 GCGGAAATCCCTCTCTGTCAAGCTGTCACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1689
QY 1741 GTGAGCAGAAAGTGGAGCTGGGGGGGTGGAGAGCCGGGCCCTTGGACCTTCCCTGACC 1800
DB 1690 GTGAGCAGAAAGTGGAGCTGGGGGGGTGGAGAGCCGGGCCCTTGGACCTTCCCTGACC 1749
QY 1801 CGTCTAATATATAAATATAGAGATGT 1826
DB 1750 CGTCTAATATATAAATATAGAGATGT 1775

RESULT 2

CR621455 1762 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0D008YK04 of Neuroblastoma Cot
DEFINITION 50-normalized of Homo sapiens (human).
ACCESSION CR621455
VERSION CR621455.1 GI:50502262
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1762)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1762)

AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	Location/Qualifiers
source	1..1762 /organism="Homo sapiens" /db_xref="taxon:9606" /cloned="CS0DD008YK04" /tissue type="Neuroblastoma Cot 50-normalized" /plasmid="pCMVSPORT_6"
ORIGIN	
	Query Match 92.5%; Score 1699.4; DB 4; Length 1762; Best Local Similarity 97.1%; Pred. No. 0; Matches 1761; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
Qy	11 GATGGCGGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAACCGAGGGGGT 70
Db	1 GATGGCGGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAACCGAGGGGGT 60
Qy	71 CGSCCCGGGGGTCCTCGGGGGAGGTGGAGATGTTGAAGGGGCAGCGGTTTCGACGTGGGCC 130
Db	61 CGSCCCGGGGGTCCTCGGGGGAGGTGGAGATGTTGAAGGGGCAGCGGTTTCGACGTGGGCC 120
Qy	131 GCCTACACGCAAGTTCAGTGATACCGCGAGGCGCGTACGGCATGCTCAGCTCGGCCCTA 190
Db	121 GCCTACACGCAAGTTCAGTGATACCGCGAGGCGCGTACGGCATGCTCAGCTCGGCCCTA 180
Qy	191 TGACCAGTGGCGAAAGCTCGCGTGGCCATCAAGAAGATCAGCCCCTTCGACATCAGAC 250
Db	181 TGACCAGTGGCGAAAGCTCGCGTGGCCATCAAGAAGATCAGCCCCTTCGAAACATCAGAC 240
Qy	251 CTACTCGACGCGACGCTCCGGAGATCCAGATCTCTGCTGCGCTTCGCCCATGAGATGT 310
Db	241 CTACTCGACGCGACGCTCCGGAGATCCAGATCTCTGCTGCGCTTCGCCCATGAGATGT 300
Qy	311 CATCGGCATCCGAGACAATTCTGCGGCGCTCCACCTCTGGAAGCCATGAGAGATGTCTACAT 370
Db	301 CATCGGCATCCGAGACAATTCTGCGGCGCTCCACCTCTGGAAGCCATGAGAGATGTCTACAT 360
Qy	371 TGTGCGAGGACCTGATGGAGATGACCTGTACAGTTGCTGAAAAGCCAGCAGCTGAGCAA 430
Db	361 TGTGCGAGGACCTGATGGAGATGACCTGTACAGTTGCTGAAAAGCCAGCAGCTGAGCAA 420
Qy	431 TGAACCATATCTGTAATCTCTTACAGATCCCTGCGGGGCTCAAGTATCATCCATCCGCG 490
Db	421 TGAACCATATCTGTAATCTCTTACAGATCCCTGCGGGGCTCAAGTATCATCCATCCGCG 480
Qy	491 CAACGTGCTCCACCGAGATCTAAAGCCCTCCAACCTGCTCATCAACACCACTCGCACCT 550
Db	481 CAACGTGCTCCACCGAGATCTAAAGCCCTCCAACCTGCTCATCAACACCACTCGCACCT 540
Qy	551 TAAGATTGTGATTTCGGCTGGCCCGGATTCGCGATCTGAGCATGACCAACCGGCTT 610
Db	541 TAAGATTGTGATTTCGGCTGGCCCGGATTCGCGATCTGAGCATGACCAACCGGCTT 600
Qy	611 CTTGAACGAGTATGTGGCTACGCGCTGGTACCGGGCCCGAGAGATCATGCTGAATCCAA 670
Db	601 CTTGAACGAGTATGTGGCTACGCGCTGGTACCGGGCCCGAGAGATCATGCTGAATCCAA 660
Qy	671 GGGCTATACCAAGTTCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTGAGATGCTCTC 730
Db	661 GGGCTATACCAAGTTCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTGAGATGCTCTC 720
Qy	731 TAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACCAATTTCTGGGCAT 790

Db 1081 AGCTCATCTTCCAGGAGACGACGCTTCAGCCGCGAGTGTGTGGAGGCCCTCCAGCCCA 1140

Qy 1156 GACAGACATCTGCAACCTGGGCTGGAAACAGAACTGCAAGAGGCAAGAGGTCACT 1215

Db 1141 GACAGACATCTGCAACCTGGGCT----- 1167

Qy 1216 GAGGGCTCTGTACCCAGGACCTGCTCTGTGCTGCTCCCTTCCCGCCAGACTGTTAGA 1275

Db 1168 -----GGACCTGCTCTGCTGCTCCCTTCCCGCCAGACTGTTAGA 1209

Qy 1276 AAATGACACTGTGCCAGCCCGACCTTGGCAGCCAGCCGCGGTGGAGCAATGGGCTT 1335

Db 1210 AAATGACACTGTGCCAGCCCGACCTTGGCAGCCAGCCGCGGTGGAGCAATGGGCTT 1269

Qy 1336 GGCACCTCTCTCTTGTGAGGCTTCCAGCTTCAAGGAGGCGACGAGGCTTCTCTCC 1395

Db 1270 GGCACCTCTCTCTTGTGAGGCTTCCAGCTTCAAGGAGGCGACGAGGCTTCTCTCC 1329

Qy 1396 CACCCGCTCTCCAGCGGCTCGGGAGCTCAGGTGGCCCAAGTTCAATCTCCCGCTGC 1455

Db 1330 CACCCGCTCTCCAGCGGCTCGGGAGCTCAGGTGGCCCAAGTTCAATCTCCCGCTGC 1389

Qy 1456 TGCTGCTGGCCCTTACCTTCCCGAGCGTCCAGTCTCTGGCAGTTCTGGAATGGAAGG 1515

Db 1390 TGCTGCTGGCCCTTACCTTCCCGAGCGTCCAGTCTCTGGCAGTTCTGGAATGGAAGG 1449

Qy 1516 TTCTGCTGCCCAACCTGTGAAGGCGAGAGGTGAGGGTGGGGGCGCTGAGTAGGA 1575

Db 1450 TTCTGCTGCCCAACCTGTGAAGGCGAGAGGTGAGGGTGGGGGCGCTGAGTAGGA 1509

Qy 1576 CTGAGGGCCATGCTGCCCGCTCATCTCATTCAAACCCACCCCTAGTTTCCCTGAAGGA 1635

Db 1510 CTGAGGGCCATGCTGCCCGCTCATCTCATTCAAACCCACCCCTAGTTTCCCTGAAGGA 1569

Qy 1636 ACATTCCTTAGTCTCAAGGCTAGCATTCCTGAGGAGCCAGGCGCGGCGAATCCCTCC 1695

Db 1570 ACATTCCTTAGTCTCAAGGCTAGCATTCCTGAGGAGCCAGGCGCGGCGAATCCCTCC 1629

Qy 1696 CTGCTAAGCTGTCATTCGCTGCGCTCGCTCTCTGCTGCTGCTGAGGAGGAGTGA 1755

Db 1630 CTGCTAAGCTGTCATTCGCTGCGCTCGCTCTCTGCTGCTGCTGAGGAGGAGTGA 1689

Qy 1756 GCTGGGGGGCGTGAGAGCCCGCGGCTCTGAGGAGCCAGGCGCGGCGAATCCCTATAA 1815

Db 1690 GCTGGGGGGCGTGAGAGCCCGCGGCTCTGAGGAGCCAGGCGCGGCGAATCCCTATAA 1749

Qy 1816 TA 1817

Db 1750 TA 1751

RESULT 5

CR597147

LOCUS

DEFINITION

full-length cDNA clone CS0D013YD03 of T cells (Jurkat cell line)

CR597147 1751 bp mRNA linear HTC 21-JUL-2004

Cot 10-normalized of Homo sapiens (human).

CR597147

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 1751)

AUTHORS

Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue

2 (bases 1 to 1751)

REFERENCE

AUTHORS

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT - Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

Location/Qualifiers

source

1..1751

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D013YD03"

/tissue_type="T cells (Jurkat cell line) Cot

10-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 91.5%; Score 1681.4; DB 4; Length 1751;

Best Local Similarity 97.1%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

Qy 1 GAGGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAA 60

Db 8 GAGGAGTGGAGATGGCGCGCGCGGCTCAGGGGGCGGGGGCGGGAGCCCCGTAGAA 67

Qy 61 CCGAGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGAGATGTTGAAGGGGAGCGGTTCG 120

Db 68 CCGAGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGAGATGTTGAAGGGGAGCGGTTCG 127

Qy 121 AGTGGGGCCGCGCTACACGAGTTGCACTACATCGCGAGGGCGGTACGGCATGTTCA 180

Db 128 AGTGGGGCCGCGCTACACGAGTTGCACTACATCGCGAGGGCGGTACGGCATGTTCA 187

Qy 181 GCTCGGCTTATGACCAACGTCGCAAGACTGCGGTGGCCATCAAGAAAGATCAGCCCTTCG 240

Db 188 GCTCGGCTTATGACCAACGTCGCAAGACTGCGGTGGCCATCAAGAAAGATCAGCCCTTCG 247

Qy 241 AACATCAGACCTTACTGCCAGCGCACTCCGGGAGATCCAGATCTCTGTCGCTTCCGCC 300

Db 248 AACATCAGACCTTACTGCCAGCGCACTCCGGGAGATCCAGATCTCTGTCGCTTCCGCC 307

Qy 301 ATGAGAAATGTCATCGCATCCGAGACATCTCTCGGGCGTCCACCTTGAAGCATGAGAG 360

Db 308 ATGAGAAATGTCATCGCATCCGAGACATCTCTCGGGCGTCCACCTTGAAGCATGAGAG 367

Qy 361 ATGTCTTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAAGCCAGC 420

Db 368 ATGTCTTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAAGCCAGC 427

Qy 421 AGCTGAGCAATGACCAATATCTCTACTTCTTACCATGATCTCTGGGGGCTTCAAGTACA 480

Db 428 AGCTGAGCAATGACCAATATCTCTACTTCTTACCATGATCTCTGGGGGCTTCAAGTACA 487

Qy 481 TCCACTCTCGCAACGCTCCACCGAGATCTAAAGCCCTCCCAACCTGCTCATCAACACCA 540

Db 488 TCCACTCTCGCAACGCTCCACCGAGATCTAAAGCCCTCCCAACCTGCTCATCAACACCA 547

Qy 541 CTTGCGACCTTAAAGATTTGTGATTTTTCGGCTTCGGCGGATTCGCCGATCTGAGCATGACC 600

Db 548 CTTGCGACCTTAAAGATTTGTGATTTTTCGGCTTCGGCGGATTCGCCGATCTGAGCATGACC 607

Qy 601 ACACGGCTTCTGACGGAGTATGCGGTACCGGTGTTACCGGGCCCCAGAGATCATGC 660

Db 608 ACACGGCTTCTGACGGAGTATGCGGTACCGGTGTTACCGGGCCCCAGAGATCATGC 667

Qy 661 TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 720

Db 668 TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 727

Qy 721 AGATGCTCTTAACCGGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACCA 780

```
Db 728 AGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 787
Qy 781 TTCTGGGCATCTGGGCTCCCATCCAGAGGACCTGTAATTCATCATCAACATGAAG 840
Db 788 TTCTGGGCATCTGGGCTCCCATCCAGAGGACCTGTAATTCATCATCAACATGAAG 847
Qy 841 CCCGAACTACTACAGTCTCTGCGCTTCAAGACCAAGGTGGCTTGGGCCAAGCTTTTC 900
Db 848 CCCGAACTACTACAGTCTCTGCGCTTCAAGACCAAGGTGGCTTGGGCCAAGCTTTTC 907
Qy 901 CCAAGTCAGATCCAAAGCCCTTGAACCTGCTGAGCCGGATGTTAACTTTAAACCCCAATA 960
Db 908 CCAAGTCAGATCCAAAGCCCTTGAACCTGCTGAGCCGGATGTTAACTTTAAACCCCAATA 967
Qy 961 AACGGATCAGGTGGAGGAGCGTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1020
Db 968 AACGGATCAGGTGGAGGAGCGTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA 1027
Qy 1021 CGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTTCGCCATGGAGCTGGATGACCTACTA 1080
Db 1028 CGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTTCGCCATGGAGCTGGATGACCTACTA 1087
Qy 1081 AGGAGCGGCTGAAGAGCTCATCTTCCAGGAGACAGACCGCTTCCAGCCCGGAGTGTGG 1140
Db 1088 AGGAGCGGCTGAAGAGCTCATCTTCCAGGAGACAGACCGCTTCCAGCCCGGAGTGTGG 1147
Qy 1141 AGGCCCTTACCCAGACAGACATCTCTGCACCTCTGGGGCTTGGACAGAACTGGCAAG 1200
Db 1148 AGGCCCTTACCCAGACAGACATCTCTGCACCTCTGGGGCTTGGACAGAACTGGCAAG 1189
Qy 1201 AGGCAAGAGTCACTGAGGGCTCTGTGCACCGAGACTTGCCTCTGCTGCCCTCTCC 1260
Db 1190 -----GGACCTGCTCTGCTGCCCTCTCC 1216
Qy 1261 CGCCAGACTGTTAGAAAATGGAACAATGTCGACCGCCGAGACTTGGAGCCGAGCCGG 1320
Db 1217 CGCCAGACTGTTAGAAAATGGAACAATGTCGACCGCCGAGACTTGGAGCCGAGCCGG 1276
Qy 1321 GTGGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTCCAGCTTCAGGCAGGCCA 1380
Db 1277 GTGGAGCATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTTCAGCTTCAGGCAGGCCA 1336
Qy 1381 AGGCTTTCTCTCCCAACCGCCCTCCGCCAGCGGCTCTGGAGCTCAGGTGGCCCACT 1440
Db 1337 AGGCTTTCTCTCCCAACCGCCCTCCGCCAGCGGCTCTGGAGCTCAGGTGGCCCACT 1396
Qy 1441 TCAATCTCCGCTGCTGCTGCGCTTACCTTCCCAAGCGTCCAGTCTCTGGCAGT 1500
Db 1397 TCAATCTCCGCTGCTGCTGCGCTTACCTTCCCAAGCGTCCAGTCTCTGGCAGT 1456
Qy 1501 TCTGGAATGGAGGGTTCTGGCTGCCCAACCTCTGAGGCGCAGAGGTGGAGGGTGGG 1560
Db 1457 TCTGGAATGGAGGGTTCTGGCTGCCCAACCTCTGAGGCGCAGAGGTGGAGGGTGGG 1516
Qy 1561 GGCGCTGAGTAGGACTCAGGGCCATGCTGCCGCCCTCATCTCATTAACACCCACCT 1620
Db 1517 GGCGCTGAGTAGGACTCAGGGCCATGCTGCCGCCCTCATCTCATTAACACCCACCT 1576
Qy 1621 AGTTTCCCTGAAGAAACATTCCTTAGTCTCAAGGGCTAGCATTCCTGAGGAGCCAGGCCG 1680
Db 1577 AGTTTCCCTGAAGAAACATTCCTTAGTCTCAAGGGCTAGCATTCCTGAGGAGCCAGGCCG 1636
Qy 1681 GGCGAATCCCTCTGCTGCAAGCTGTCACTTGGCTGCCCTGCTGCTTCTGTGTG 1740
Db 1637 GGCGAATCCCTCTGCTGCAAGCTGTCACTTGGCTGCCCTGCTGCTTCTGTGTG 1696
Qy 1741 GTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCCCGGCCCTTGGCACCTCC 1795
Db 1697 GTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCCCGGCCCTTGGCACCTCC 1751
```

RESULT 6

```
CR603463 1854 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DN005YA14 of Adult brain of Homo sapiens
(human)
DEFINITION
ACCESSION CR603463
VERSION 1 GI:50484270
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1854)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1854)
Direct Submission
Genoscope.
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1854
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN005YA14"
/tissue_type="Adult brain"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 91.1%; Score 1673.4; DB 4; Length 1854;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1762; Conservative 0; Mismatches 1; Indels 58; Gaps 2;
Qy 1 GAGGAGTGGAGATGGCGGCGCGGCTCAGGGGCGGGGCGGGAGCCCGGTAGAA 60
Db 75 GAGGAGTGGAGATGGCGGCGCGGCTCAGGGGCGGGGCGGGAGCCCGGTAGAA 134
Qy 61 CCAGAGGGGTTCGGCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTCC 120
Db 135 CCAGAGGGGTTCGGCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGAGCCGTTCC 194
Qy 121 ACCTGGGCGCGCTACACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 180
Db 195 ACCTGGGCGCGCTACACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 254
Qy 181 GCTCGGCTTATGACCAAGTTCGCAAGTTCGCTGGCCATCAAGAGATCAGCCCTTCG 240
Db 255 GCTCGGCTTATGACCAAGTTCGCAAGTTCGCTGGCCATCAAGAGATCAGCCCTTCG 314
Qy 241 AACATCAGACTTACTGCGCAGCGCTCCGGGAGATCCAGATCTCTGCGTTCGCGCC 300
Db 315 AACATCAGACTTACTGCGCAGCGCTCCGGGAGATCCAGATCTCTGCGTTCGCGCC 374
Qy 301 ATGAGATGTCATCGGCTCCGAGACATTCCTGGGGGCTCCACCCTGGAAGCCATGAGAG 360
Db 375 ATGAGATGTCATCGGCTCCGAGACATTCCTGGGGGCTCCACCCTGGAAGCCATGAGAG 434
Qy 361 ATGTCTACATTTGTGAGGACCTGATGGAGATGACCTGTGTAAGTTCTGAAAGCCAGC 420
Db 435 ATGTCTACATTTGTGAGGACCTGATGGAGATGACCTGTGTAAGTTCTGAAAGCCAGC 494
Qy 421 AGCTGAGCAATGACCATATCTGCTACTTCTTCTTACCAGATCTTGGGGGCTCAAGTACA 480
```


495	Db		AGCTGAGCAATGACCAATATCTGCTACTTCTCTACCAAGATCCTGCGGGGCTCAAGTACA	554
481	Qy		TCCACTCCGCCCAAGTGTCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA	540
555	Db		TCCACTCCGCCCAAGTGTCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA	614
541	Qy		CCTGCGACCTTAAGATTTGTGATTTCCGCGCTGGCCCGGATTTGCCGATCTCTGAGCATGACC	600
615	Db		CCTGCGACCTTAAGATTTGTGATTTCCGCGCTGGCCCGGATTTGCCGATCTCTGAGCATGACC	674
601	Qy		ACACCGGCTTCTTGACCGAGTATGTGCTTACGGCTGCGCTGCTACCGGGCCCCAGAGATCATGC	660
675	Db		ACACCGGCTTCTTGACCGAGTATGTGCTTACGGCTGCGCTGCTACCGGGCCCCAGAGATCATGC	734
661	Qy		TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG	720
735	Db		TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG	794
721	Qy		AGATGCTCTTAACCGGGCCCATCTTCCCTGGCAAGCACTACTCTGGATCAGCTCAACCCACA	780
795	Db		AGATGCTCTTAACCGGGCCCATCTTCCCTGGCAAGCACTACTCTGGATCAGCTCAACCCACA	854
781	Qy		TTCTGGGCTCATCTGGGCTCCCAATCCAGGAGGACTGAATTTGATCATCAACATGAAGG	840
855	Db		TTCTGGGCTCATCTGGGCTCCCAATCCAGGAGGACTGAATTTGATCATCAACATGAAGG	914
841	Qy		CCCGAAATCTACATCAGTCTCTGCCCTCCAAGCAAGGTGGCTTGGGCCCAAGCTTTTCC	900
915	Db		CCCGAAATCTACATCAGTCTCTGCCCTCCAAGCAAGGTGGCTTGGGCCCAAGCTTTTCC	974
901	Qy		CCAACTCAGACTCCAAA-----GCCCTTGACCTGCTGGACCGGATGTT	943
975	Db		CCAACTCAGACTCCAAAGGTTGAGAGACATGAGGCCCTTGACCTGCTGGACCGGATGTT	1034
944	Qy		AACCTTTAAACCCCAATAAACGAGTCAAGTGGAGGAAGCGTGGCTCACCCCTACCTGGA	1003
1035	Db		AACCTTTAAACCCCAATAAACGAGTCAAGTGGAGGAAGCGTGGCTCACCCCTACCTGGA	1094
1004	Qy		GCAGTACTATGACCCGACGGATGAGCCAGTGGCCGAGGAGGCCCTTCACTTCGCGCATGGA	1063
1095	Db		GCAGTACTATGACCCGACGGATGAGCCAGTGGCCGAGGAGGCCCTTCACTTCGCGCATGGA	1154
1064	Qy		GCTGGATGACCTACTTAAGGAGCGGCTGAAGGAGCTCATCTTCCAGAGAGACAGCACGCTT	1123
1155	Db		GCTGGATGACCTACTTAAGGAGCGGCTGAAGGAGCTCATCTTCCAGAGAGACAGCACGCTT	1214
1124	Qy		CCAGCCCGGAGTGTGTGAGGCCCCCTTAGCCCAGACAGACATCTCTGCACTCTGGGGCTG	1183
1215	Db		CCAGCCCGGAGTGTGTGAGGCCCCCTTAGCCCAGACAGACATCTCTGCACTCTGGGGCT-	1273
1184	Qy		GAA CAGAACTGGCAAGAGGCAAGAGTCACTGAGGGGCTCTGTACCCAGGAGCTGCTGCT	1243
1274	Db		-----GGACCTCGCTT-----GGACCTCGCTT	1283
1244	Qy		CCTGCTGCCCTCTCCCGCCAGACTGTTAGAAAATGAGACACTGTGCCACAGCCCGGACCT	1303
1284	Db		CCTGCTGCCCTCTCCCGCCAGACTGTTAGAAAATGAGACACTGTGCCACAGCCCGGACCT	1343
1304	Qy		TGGCAGCCCAAGCCCGGGGTGGAGCATGGGCTCGGCCACTCTCTCTCTTTGCTGAGGCTC	1363
1344	Db		TGGCAGCCCAAGCCCGGGGTGGAGCATGGGCTCGGCCACTCTCTCTCTTTGCTGAGGCTC	1403
1364	Qy		CAGCTTCAGGAGGCGCAAGGCTTCTCTCTCCCAACCGGCCCTCCCAAGGGGCTCGGGA	1423
1404	Db		CAGCTTCAGGAGGCGCAAGGCTTCTCTCTCCCAACCGGCCCTCCCAAGGGGCTCGGGA	1463
1424	Qy		GCTCAGGTGGCCCAAGTTCATCTCCCGCTGCTGCTGGCGCTTACCTTCCCAAGCG	1483
1464	Db		CCTCAGGTGGCCCAAGTTCATCTCCCGCTGCTGCTGGCGCTTACCTTCCCAAGCG	1523
1484	Qy		TCCAGTCTCTGGCAGTTCGGAATGGAAGGTTCTGCTGCTGCCCAACCTGCTGAAGGC	1543

Db	1524	TCC	CAGTCTCTGGCAGGTTCTGGAATGAAGGGTTCTGGCTGCCCAACCTGCTGAAGGGC	1584
Qy	1544	AGAGGTGGAGGGTGGGGGGCGCTGAGTAGGGA	CTCAGGGCCATGCTGCCGCCCTCATCT	1603
Db	1584	AGAGGTGGAGGGTGGGGGGCGCTGAGTAGGGA	CTCAGGGCCATGCTGCCGCCCTCATCT	1643
Qy	1604	CATTCAAAACCCCAACCTTAGTTCCTTGAAGGA	CATTTCCTTAGTCTCAAGGGCTAGCATC	1663
Db	1644	CATTCAAAACCCCAACCTTAGTTCCTTGAAGGA	CATTTCCTTAGTCTCAAGGGCTAGCATC	1703
Qy	1664	CCTGAGAGCCAGCGCGGGCGGAATCCCTCCCTGTC	CAAGAGCTGCTCACTTCGGGTGCCCT	1723
Db	1704	CCTGAGAGCCAGCGCGGGCGGAATCCCTCCCTGTC	CAAGAGCTGCTCACTTCGGGTGCCCT	1763
Qy	1724	CGCTGCTTCTGTGTGCTGAGCAGAAGTGGAGCT	TGGGGCGCTGGAGAGCCCGGCGCCC	1783
Db	1764	CGCTGCTTCTGTGTGCTGAGCAGAAGTGGAGCT	TGGGGCGCTGGAGAGCCCGGCGCCC	1823
Qy	1784	CTGCCACCTCCCTGACCCCGTCTAATATATAA	1814	
Db	1824	CTGCCACCTCCCTGACCCCGTCTAATATATAA	1854	
RESULT 7				
CR621977				
LOCUS				
DEFINITION				
full-length cDNA clone CS0DJ015YL15 of T cells (Jurkat cell line)				
Cot 10-normalized of Homo sapiens (human).				
ACCESSION				
CR621977				
VERSION				
CR621977.1 GI:50502784				
KEYWORDS				
HTC; CNSUT_CDNA.				
SOURCE				
Homo sapiens				
ORGANISM				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
Hominidae; Homo.				
REFERENCE				
1 (bases 1 to 1729)				
AUTHORS				
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.				
TITLE				
Full-length cDNA libraries and normalization				
JOURNAL				
Unpublished				
REMARK				
Contact : Feng Liang Email : fliang@lifetech.com URL :				
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				
Faraday Avenue				
2 (bases 1 to 1729)				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
Submitted (20-JUL-2004) Genoscope - Centre National de Sequences ;				
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
- Web : www.genoscope.cns.fr)				
COMMENT				
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer				
end enriched, double-strand cDNA was digested with Not I and cloned				
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library				
was normalized. Library was constructed by Life Technologies, a				
division of Invitrogen.				
FEATURES				
source				
1. .1729				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="CS0DJ015YL15"				
/tissue_type="T cells (Jurkat cell line) Cot				
10-normalized"				
/plasmid="pCMVSPORT_6"				
ORIGIN				
Query Match 90.7%; Score 1666.4; DB 4; Length 1729;				
Best Local Similarity 97.1%; Pred. No. 0;				
Matches 1728; Conservative 0; Mismatches 1; Indels 51; Gaps 1;				
Qy	36	GGCGGGGGCGGGAGACCCCGCTAGAACCAAGGGGTTCGCCCGGGGTCCCGGGGAGGTG	95	
pb	1	GGCGGGGGCGGGAGACCCCGCTAGAACCAAGGGGTTCGCCCGGGGTCCCGGGGAGGTG	60	

QY 96 GAGATGGTGAAGGGGCGAGCGTTTCGACGTGGGCGCGCTTACACGACAGTTGCGAGTACATC 155
DB |||||
61 GAGATGGTGAAGGGGCGAGCGTTTCGACGTGGGCGCGCTTACACGACAGTTGCGAGTACATC 120
QY 156 GCGAGGGGCGGTACGGCATGGTTCAGCTCGGCGCTATGACACAGTGGCGAGACTCCGGTG 215
DB |||||
121 GCGAGGGGCGGTACGGCATGGTTCAGCTCGGCGCTATGACACAGTGGCGAGACTCCGGTG 180
QY 216 GCCATCAAGAAGATCAGACGCCCTTCGAACATCAGACCTTACTGCGCAGGCGACGCTCCGGGAG 275
DB |||||
181 GCCATCAAGAAGATCAGACGCCCTTCGAACATCAGACCTTACTGCGCAGGCGACGCTCCGGGAG 240
QY 276 ATCCAGATCTGCTGGCGCTTCGCGCATAGAGATGTGATCGGCAATCCGAGACATCTCGCG 335
DB |||||
241 ATCCAGATCTGCTGGCGCTTCGCGCATAGAGATGTGATCGGCAATCCGAGACATCTCGCG 300
QY 336 GCGTCCACCTGGAAGCCATGAGAGATGTCTACATTTGTCAGGACCTGATGGAGACTGAC 395
DB |||||
301 GCGTCCACCTGGAAGCCATGAGAGATGTCTACATTTGTCAGGACCTGATGGAGACTGAC 360
QY 396 CTGTACAAAGTTGCTGAAAGCCAGCAGCTGAGCAATGACCATATCTGTACTTCTCTTAC 455
DB |||||
361 CTGTACAAAGTTGCTGAAAGCCAGCAGCTGAGCAATGACCATATCTGTACTTCTCTTAC 420
QY 456 CAGATCTTCGCGGGCCTCAAGTATCATCACTCCGCCAACTGTCTCAACCGAGATCTAAAG 515
DB |||||
421 CAGATCTTCGCGGGCCTCAAGTATCATCACTCCGCCAACTGTCTCAACCGAGATCTAAAG 480
QY 516 CCCTCCAACTGTCTATCAACACCTGCGACCTTAAGATTTGTGATTTTCGGCCTGGCC 575
DB |||||
481 CCCTCCAACTGTCTATCAACACCTGCGACCTTAAGATTTGTGATTTTCGGCCTGGCC 540
QY 576 CGGATTTGCCGATCTGAGCATGACCAACCGGCTTCTGACGAGTATGTGGCTAGCGCG 635
DB |||||
541 CGGATTTGCCGATCTGAGCATGACCAACCGGCTTCTGACGAGTATGTGGCTAGCGCG 600
QY 636 TGGTACCGGGCCCCAGAGATCATCTGAACTTCAAGGGCTATACCAAGTCCATCGAATC 695
DB |||||
601 TGGTACCGGGCCCCAGAGATCATCTGAACTTCAAGGGCTATACCAAGTCCATCGAATC 660
QY 696 TGGTCTGTGGGCTGCATTCTGGCTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAG 755
DB |||||
661 TGGTCTGTGGGCTGCATTCTGGCTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAG 720
QY 756 CACTTACCTGATCAGCTCAACCACTTCTGGGATCTCTGGGCTCCCGCATCCAGGAGGAC 815
DB |||||
721 CACTTACCTGATCAGCTCAACCACTTCTGGGATCTCTGGGCTCCCGCATCCAGGAGGAC 780
QY 816 CTGAATTTGATCATCAACATGAAGGCCCGAACTTACCTACAGTCTCTGCGCTCCAGACC 875
DB |||||
781 CTGAATTTGATCATCAACATGAAGGCCCGAACTTACCTACAGTCTCTGCGCTCCAGACC 840
QY 876 AAGTGGCTTGGGCAAGCTTTTCCCAAGTCAAGTCCAAAGCCCTTGACCTGTGGAC 935
DB |||||
841 AAGTGGCTTGGGCAAGCTTTTCCCAAGTCAAGTCCAAAGCCCTTGACCTGTGGAC 900
QY 936 CGGATTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTT 995
DB |||||
901 CGGATTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTT 960
QY 996 TACCTGGAGCAGTACTATGACCCGACGAGTGGCCAGTGGCGGAGGAGCCCTTCACTTC 1055
DB |||||
961 TACCTGGAGCAGTACTATGACCCGACGAGTGGCCAGTGGCGGAGGAGCCCTTCACTTC 1020
QY 1056 GCCATGAGCTGGATGATCTTACCTTAAGAGCGGCTGAAGAGGCTCATCTTCCAGGAGACA 1115
DB |||||
1021 GCCATGAGCTGGATGATCTTACCTTAAGAGCGGCTGAAGAGGCTCATCTTCCAGGAGACA 1080
QY 1116 GCAGCTTCCAGCCGCGAGTGTCTGGAGGCCCTTAGCCGACAGACATCTCTGCACTTC 1175
DB |||||
1081 GCAGCTTCCAGCCGCGAGTGTCTGGAGGCCCTTAGCCGACAGACATCTCTGCACTTC 1140
QY 1176 GGGGCTTGGAAAGAACTGGCAAGAGGCAAGAGGTCACTGAGGGCCTCTGTCAACCCAGG 1235

DB 1141 GGGGCT-----GG 1149
QY 1236 AACTGCTCTCTGCTGCCCCCTCTCCCGCAGACTGTTAGAAAATGGACACTGTGCCAGC 1295
DB |||||
1150 AACTGCTCTCTGCTGCCCCCTCTCCCGCAGACTGTTAGAAAATGGACACTGTGCCAGC 1209
QY 1296 CCGGACCTTGGCAGCCCGAGCGGGGTGAGCATGGCCCTGSCCACTCTCTCTTTGCT 1355
DB |||||
1210 CCGGACCTTGGCAGCCCGAGCGGGGTGAGCATGGCCCTGSCCACTCTCTCTTTGCT 1269
QY 1356 GAGGCTCTCAGCTTCAAGGAGCGCAAGGCTTCTCTCTCCCGCCCTCCCGCAGCGGG 1415
DB |||||
1270 GAGGCTCTCAGCTTCAAGGAGCGCAAGGCTTCTCTCTCCCGCCCTCCCGCAGCGGG 1329
QY 1416 CTTGCGGAGCTCAGGTGGCCCCCAGTTCAATCTCCGCTGCTGTGCTGGCCCTTACCTT 1475
DB |||||
1330 CTTGCGGAGCTCAGGTGGCCCCCAGTTCAATCTCCGCTGCTGTGCTGGCCCTTACCTT 1389
QY 1476 CCGCAGCTCCAGTCTCTGGCAGTTCTGGAATGGAAGGTTCTGGCTGCCCAACTGC 1535
DB |||||
1390 CCGCAGCTCCAGTCTCTGGCAGTTCTGGAATGGAAGGTTCTGGCTGCCCAACTGC 1449
QY 1536 TGAAGGCGAGAGTGGAGGGTGGGGGCGCTGAGTAGGAGACTCAGGGCCATGCTGCCCC 1595
DB |||||
1450 TGAAGGCGAGAGTGGAGGGTGGGGGCGCTGAGTAGGAGACTCAGGGCCATGCTGCCCC 1509
QY 1596 CTTCACTCTCATTAACACCCACCTAGTTTCCCTGAAGGAACAATTCCTTAGTCTCAAGGG 1655
DB |||||
1510 CTTCACTCTCATTAACACCCACCTAGTTTCCCTGAAGGAACAATTCCTTAGTCTCAAGGG 1569
QY 1656 CTAGCATTCCTGAGGAGCAGCGCGGCGGAATCCCTCTCCCTGTCAAGTGTCACTTCG 1715
DB |||||
1570 CTAGCATTCCTGAGGAGCAGCGCGGCGGAATCCCTCTCCCTGTCAAGTGTCACTTCG 1629
QY 1716 CGTGGCCTCGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1775
DB |||||
1630 CGTGGCCTCGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1689
QY 1776 CGCGCGCCCTGCCACCTCCCTGACCCCTCTAATATATATAA 1815
DB |||||
1690 CGCGCGCCCTGCCACCTCCCTGACCCCTCTAATATATATAA 1729

RESULT 8

CR621688
LOCUS full-length cDNA clone CS0D1044VF16 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).

CR621688

ACCESSION CR621688.1 GI:50502495

VERSION HTC; CNSLT cdNA.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1707)

AUTHORS Li, W. B., Gruber, C., Jesse, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

REMARK Unpublished

CONTACT Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue

REFERENCE 2 (bases 1 to 1707)

AUTHORS Genoscope.

TITLE Direct Submission

REMARK Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

ORGANISM	Homo sapiens	Db
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1678)	
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE	Full-length cDNA libraries and normalization	
JOURNAL	Unpublished	
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue	
REFERENCE	2 (bases 1 to 1678)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)	
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.	
FEATURES	Location/Qualifiers	
source	1..1678	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="CS0DC016YM24"	
	/tissue_type="Neuroblastoma Cot 25-normalized"	
	/plasmid="pCMVSPORT_6"	
ORIGIN		
Query Match	87.9%; Score 1615.4; DB 4; Length 1678;	
Best Local Similarity	97.0%; Pred. No. 0;	
Matches 1677; Conservative	0; Mismatches 1; Indels 51; Gaps 1;	
Qy	73 GCCCGGGGTCCTCGGGGAGGTGGAGATGTGTGAAGGGGACGCGTTTCGAGTGGGCGCGC 132	
Db	1 GCCCGGGGTCCTCGGGGAGGTGGAGATGTGTGAAGGGGACGCGTTTCGAGTGGGCGCGC 60	
Qy	133 GCTACACGAGTTTCAGTACATCGGGGAGCGGTACGGCATGGTTCAGCTCGGCGCTATG 192	
Db	61 GCTACACGAGTTTCAGTACATCGGGGAGCGGTACGGCATGGTTCAGCTCGGCGCTATG 120	
Qy	193 ACCACGTGCGCAAGACTCGCGTGGCCATCAAGAGATCAGCCCTTCGAAACATCAGACCT 252	
Db	121 ACCACGTGCGCAAGACTCGCGTGGCCATCAAGAGATCAGCCCTTCGAAACATCAGACCT 180	
Qy	253 ACTGCCAGCGCAGCTCCGGGAGATCCAGATTCCTGTGCGCTTCGCCCATGAGAAATGCA 312	
Db	181 ACTGCCAGCGCAGCTCCGGGAGATCCAGATTCCTGTGCGCTTCGCCCATGAGAAATGCA 240	
Qy	313 TCGGCATCCGAGACATTCGGGGGCGTCCACCTTGAAGCCATGAGAGATGCTCACAATTG 372	
Db	241 TCGGCATCCGAGACATTCGGGGGCGTCCACCTTGAAGCCATGAGAGATGCTCACAATTG 300	
Qy	373 TGCAGGACCTGATGGAGACTGACCTGTACAGTTGTCTGAAAGCCAGCAGCTGAGCAATG 432	
Db	301 TGCAGGACCTGATGGAGACTGACCTGTACAGTTGTCTGAAAGCCAGCAGCTGAGCAATG 360	
Qy	433 ACCATATCTGTACTTCTCTACAGATTCCTGCGGGGCTCAAGTACATCCACTCCGCCA 492	
Db	361 ACCATATCTGTACTTCTCTACAGATTCCTGCGGGGCTCAAGTACATCCACTCCGCCA 420	
Qy	493 ACGTGTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACCTCGGACCTTA 552	
Db	421 ACGTGTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACCTCGGACCTTA 480	
Qy	553 AGATTGTGATTTTCGGCTCGCCGGATTGCGGATTCCTGAGCATGACACACCGGCTTCC 612	
Db	481 AGATTGTGATTTTCGGCTCGCCGGATTGCGGATTCCTGAGCATGACACACCGGCTTCC 540	
Qy	613 TGACGGAGTATGTGGCTACGCGTGGTACCGGGGCCCCAGAGATCATGCTGAACTCCAAGG 672	
541	TGACGGAGTATGTGGCTACGCGTGGTACCGGGGCCCCAGAGATCATGCTGAACTCCAAGG 600	
673	GCTATACCAAGTCCATCGACATCTGCTGTGTGGGCTGCATTCGTGGCTGAGATGCTCTCTA 732	
601	GCTATACCAAGTCCATCGACATCTGCTGTGTGGGCTGCATTCGTGGCTGAGATGCTCTCTA 660	
733	ACGGGCCCATCTTCCTCGCAAGCACTACTGTGATCAGTCAACACATTCCTGGGATCC 792	
661	ACGGGCCCATCTTCCTCGCAAGCACTACTGTGATCAGTCAACACATTCCTGGGATCC 720	
793	TGGGCTCCCCATCCAGGAGGACCTGAATTGTATCATCAACATGAAGGCCCGAAATACC 852	
721	TGGGCTCCCCATCCAGGAGGACCTGAATTGTATCATCAACATGAAGGCCCGAAATACC 780	
853	TACAGTCTCTGCGCTCCAAAGCAAGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGACT 912	
781	TACAGTCTCTGCGCTCCAAAGCAAGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGACT 840	
913	CMAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAACCCCAATTAACGGATCACAG 972	
841	CMAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAACCCCAATTAACGGATCACAG 900	
973	TGGAGGAAGCGCTGGCTCACCCCTACTCGAGCAGTACTATGACCCGACGATGAGCCAG 1032	
901	TGGAGGAAGCGCTGGCTCACCCCTACTCGAGCAGTACTATGACCCGACGATGAGCCAG 960	
1033	TGGCCGAGGAGCCCTTACCTTCGCCATCGAGCTGGATGACCTTACTTAAGGAGCGGCTGA 1092	
961	TGGCCGAGGAGCCCTTACCTTCGCCATCGAGCTGGATGACCTTACTTAAGGAGCGGCTGA 1020	
1093	AGAGCTCATCTTCAGGAGACAGCAGCTTCAGCCCGGAGTGTGGAGGCCCCCTTACC 1152	
1021	AGAGCTCATCTTCAGGAGACAGCAGCTTCAGCCCGGAGTGTGGAGGCCCCCTTACC 1080	
1153	CCAGACAGACATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAAAGAGCAAGGTC 1212	
1081	CCAGACAGACATCTCTGCACCCCTGGGGCTTGGAAACAGAACTGGCAAAAGAGCAAGGTC 1110	
1213	ACTGAGGGGCTCTGTCTCACCCAGGACCTGCTCTCTGCTGCTCCCTTCTCCCGCCAGACTGTT 1272	
1111	-----GGACCTGCTCTCTGCTGCTCCCTTCTCCCGCCAGACTGTT 1149	
1273	AGAAATGGAACATGTGTCCCGACCGCCGACCTTGGACCCAGCCCGGGTGGAGCATGGG 1332	
1150	AGAAATGGAACATGTGTCCCGACCGCCGACCTTGGACCCAGCCCGGGTGGAGCATGGG 1209	
1333	CCTGGCCACCTCTCTCTTTGCTGAGGCTTCCAGCTTTCAGGAGGCCAAGCCCTTCTCCT 1392	
1210	CCTGGCCACCTCTCTCTTTGCTGAGGCTTCCAGCTTTCAGGAGGCCAAGCCCTTCTCCT 1269	
1393	CCCAACCGCCCTCCCAACCGGGGCTCGGGAGCTCAGGTGGCCCCCAAGTTCAATCTCCCGC 1452	
1270	CCCAACCGCCCTCCCAACCGGGGCTCGGGAGCTCAGGTGGCCCCCAAGTTCAATCTCCCGC 1329	
1453	TGCTGTCTGCGCCCTTACCTTCCCGACGCTCCAGTCTCCAGTCTCTGGCAGTCTCGAATGGA 1512	
1330	TGCTGTCTGCGCCCTTACCTTCCCGACGCTCCAGTCTCTGGCAGTCTCGAATGGA 1389	
1513	GGGTTCGTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGGGGCGCTGAGTAG 1572	
1390	GGGTTCGTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGGGGCGCTGAGTAG 1449	
1573	GGACTCAGGGCCATGCTGCCCTCCCTCATCTCATTTCAAAACCCCAAGTTCCTTCCCTGAA 1632	
1450	GGACTCAGGGCCATGCTGCCCTCCCTCATCTCATTTCAAAACCCCAAGTTCCTTCCCTGAA 1509	
1633	GGAAACATTTCTTGTAGTCTCAAGGGGTAGCATCCCTGAGGAGCCAGGCGCGGATCCCC 1692	
1510	GGAAACATTTCTTGTAGTCTCAAGGGGTAGCATCCCTGAGGAGCCAGGCGCGGATCCCC 1569	
1693	TCCCTGTCAAGCTGTACATTCGGTGGCCCTCGCTGCTCTCTGTGTGTGTGAGCAGAGT 1752	


```

QY 855 CAGTCTCTGCCCTCCAAAGCAAAAGGTGGCTTGGGCCAAAGCTTTTCCCAAAGTCAGACTCC 914
DB 844 CAGTCTCTGCCCTCCAAAGCAAAAGGTGGCTTGGGCCAAAGCTTTTCCCAAAGTCAGACTCC 903
QY 915 AAAGCCCTTACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAACCAGATCAGATG 974
DB 904 AAAGCCCTTACCTGCTGGACCGGATGTTAACTTTAAACCCCAATAAACCAGATCAGATG 963
QY 975 GAGGAAGCGTGGCTCACCCCTTACCTGGAGCAGTACTATGACCCGACCGAGTACCCAGTG 1034
DB 964 GAGGAGCGTGGCTCACCCCTTACCTGGAGCAGTACTATGACCCGAGGATGATCCAGTT 1023
QY 1035 GCGGAGGAGCGCTTACCTTGGCCATGGAGCTGATGACCTACTAAGGAGCGGCTGAAG 1094
DB 1024 GSCGAGGAGCGCTTCAACGCT--CGCATGGAGCTGATGATGATGATGATGATGATGATG 1081
QY 1095 GAGCT 1099
DB 1082 YATCT 1086

RESULT 13
AL530654/c
LOCUS AL530654 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DD008YK04 3-PRIME, mRNA sequence.
ACCESSION AL530654
VERSION AL530654.2 GI:31068487
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:12794147.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8342.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?n=CS0DD008BF02NFI&c=8342.f.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD008YK04"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 54.9%; Score 1008; DB 1; Length 1201;
Best Local Similarity 90.6%; Pred. No. 4.7e-218;
Matches 1140; Conservative 36; Mismatches 23; Indels 59; Gaps 8;

QY 568 GCCTGGCCGAGTTCGGATCTGAGCATGACACACCGGCTTCTGACGAGTATGTGG 627
DB 1201 GCCTTGGCCGAGTTCGGATCTGAGCATGACACACCGGCTTCTGACGAGTATGTGG 1144

```

```

QY 628 CTAAGCGCTGGTACCGGGCCCGAGAGATCATGCTGAACCTCCAGGSGCTATACCAAGTCCA 687
DB 1143 C-ACGGGBTGTACACCGGSCCMAGAGTTCAGGCTGAACCTCCAGGGYTTATACCAAGTCCA 1085
QY 688 TCGACATCTGGTCTGTGGGCTGCATTTCTGGCTCAGATGCTCTTAACCGGCCCCATCTTCC 747
DB 1084 TCGACATTTGTTCTGTGGGCTGMAATTTCTGGCTAAGATGTTCTTAACCGGCCCCWTTT-C 1026
QY 748 CTGGCAAGCACTACCTGGATCAGCTCAACCAACATTTCTGGGCACTCTGGGGC-TCGCCATCC 806
DB 1025 CTTGMAAGCACTACCTGGATCAGCTCAACCAACATTTCTGGGCACTCTGGGGCTCCCCATCC 966
QY 807 CAGGAGGACCTGATTTGATATCATCAACATGAAGGCCCGGAACCTACCTACAGCTCTGGCC 866
DB 965 CAGGAGGACCTGATTTGATATCATCAACATGAAGGCCCGGAACCTACCTACAGCTCTGGCC 906
QY 867 TCCAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGACTCCAAAGGCCCTTGAC 926
DB 905 TCCAAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGACTCCAAAGGCCCTTGAC 846
QY 927 CTGCTGGACCGGATGTTAACTTTAAACCCCAATAAACCAGATCAGTGGAGGAAGCGCTG 986
DB 845 CTGCTGGACCGGATGTTAACTTTAAACCCCAATAAACCAGATCAGTGGAGGAAGCGCTR 786
QY 987 GCTACACCTTACCTGGAGCAGTACTATGACCCGACCGATGAGCCAGTGGCCGAGGAGCCC 1046
DB 786 GCTACACCTTACCTGGAGCAGTACTATGACCCGACCGATGAGCCAGTGGCCGAGGAGCCC 726
QY 1047 TTCACCTTTCGCCATGGAGCTGGATGACCTTAAAGAGCGGCTGAAGGAGCTCATCTTC 1106
DB 725 TTCACCTTTCGCCATGGAGCTGGATGACCTTAAAGAGCGGCTGAAGGAGCTCATCTTC 666
QY 1107 CAGGAGACGACGCTTTCCAGCCCGGAGTGTCTGGAGCCGCCCTTAGCCACAGACATCT 1166
DB 665 CAGGAGACGACGCTTTCCAGCCCGGAGTGTCTGGAGCCGCCCTTAGCCACAGACATCT 606
QY 1167 CTGACCTTGGGGCTTGGAAACAGAACTGGCAAGAGCAAGAGTCACTGAGGSGCTCTG 1226
DB 605 CTGACCTTGGGGCTTGGAAACAGAACTGGCAAGAGCAAGAGTCACTGAGGSGCTCTG 590
QY 1227 TCACCCAGGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286
DB 589 -----GGACCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 537
QY 1287 GTGCCACAGCCCGGACCTTGGCAGCCAGCCGCGGGGTGGAGCATGGGCTTGGCCACCTCTC 1346
DB 536 GTGCCACAGCCCGGACCTTGGCAGCCAGCCGCGGGGTGGAGCATGGGCTTGGCCACCTCTC 477
QY 1347 TCTTTTGTGAGGCTTCCAGCTTTCAGGCGAGGCCAAGGCTTCTCTCCCAACCGGCTCTC 1406
DB 476 TCTTTTGTGAGGCTTCCAGCTTTCAGGCGAGGCCAAGGCTTCTCTCCCAACCGGCTCTC 417
QY 1407 CCAACCGGCTTGGGAGCTCAGGTGGCCCGGCTTCAATCTCCGCTGCTGCTGCTGCTGCTG 1466
DB 416 CCAACCGGCTTGGGAGCTCAGGTGGCCCGGCTTCAATCTCCGCTGCTGCTGCTGCTGCTG 358
QY 1467 CTTTACCTTCCCAAGCGCTTCCAGCTTCTGGCAGTTCCTGGAATGGAAGGGTCTTGGCTGCC 1526
DB 357 CTTTACCTTCCCAAGCGCTTCCAGCTTCTGGCAGTTCCTGGAATGGAAGGGTCTTGGCTGCC 298
QY 1527 CCAACCTTGTGAAGGGCAGAGGTGGAGGGTGGGGGGCGCTGAGTGGAGCATCAGGCGCAT 1586
DB 297 CCAACCTTGTGAAGGGCAGAGGTGGAGGGTGGGGGGCGCTGAGTGGAGCATCAGGCGCAT 238
QY 1587 GCCTGGCCCCCTCATCTCATTTCAAAACCCCAACCTAGTTTCCCTGGAAGGAACATTCCTTAG 1646
DB 237 GCCTGGCCCCCTCATCTCATTTCAAAACCCCAACCTAGTTTCCCTGGAAGGAACATTCCTTAG 178
QY 1647 TCTCAAGGGCTAGCATCCCTGAGGAGCCAGCGCGGGCGGGAATCCCTCTCCCTGTCAAGCT 1706
DB 177 TCTCAAGGGCTAGCATCCCTGAGGAGCCAGCGCGGGCGGGAATCCCTCTCCCTGTCAAGCT 118

```



```
QY 1707 GTCACTTCGGCTGCCCTCGCTGCTTCTGTGTGTGGTGAGCAGAGT -GGAGCTGGGGGC 1765
|||||
Db 117 GTCACTTCGGCTGCCCTCGCTGCTTCTGTGTGTGGTGAGCAGAGTGGGGGC 58
|||||

QY 1766 GTGGAGAGCCCGCCCTGCGACCTCCCTGACCCCTCTAATATATAATATAGAGA 1823
|||||
Db 57 GTGGAGAGCCC -GMACCCTGCCRCCTSCGTGACCCCGTCTAAWATATAATATAGAGA 1

RESULT 14
AL530980 1016 bp mRNA linear EST 24-MAR-2004
LOCUS AL530980 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DD003YH16 5-PRIME, mRNA sequence.
ACCESSION AL530980
VERSION AL530980.3 GI:45705927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1016)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31069813.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8342.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DD003DD08QP1&c=8342.f.
FEATURES
source
1. .1016
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD003YH16"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 52.3%; Score 960.2; DB 1; Length 1016;
Best Local Similarity 99.2%; Pred. No. 3.3e-207;
Matches 971; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 1 GAGGAGTGAGAGATGGCGCGCGCGCTCAGGGGGCGGGGGGAGCCCCCGTAGAA 60
|||||
Db 7 GAGGAGTGAGAGATGGCGCGCGCGCTCAGGGGGCGGGGGGAGCCCCCGTAGAA 66
|||||

QY 61 CCAGAGGGGTCCGCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACCGCTCG 120
|||||
Db 67 CCAGAGGGGTCCGCGCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACCGCTCG 126
|||||

QY 121 ACCTGGGCGCGCGCTACACGAGTTGAGTACATCGCGAGGGCGGGCTAGCGATGTCA 180
|||||
Db 127 ACCTGGGCGCGCTACACGAGTTGAGTACATCGCGAGGGCGGGCTAGCGATGTCA 186
|||||

QY 181 GCTCGGCTATGACGACGTGGCAAGACTCGCTGGCCATCAAGAGATCAGCCCCCTTCG 240
|||||
Db 187 GCTCGGCTATGACGACGTGGCAAGACTCGCTGGCCATCAAGAGATCAGCCCCCTTCG 246
|||||
```

```
QY 241 AACATCAGACCTACTGCGAGCGCAGCTCCGGGAGATCCAGATCTCTGCTGCGCTTCCGCC 300
|||||
Db 247 AACATCAGACCTACTGCGAGCGCAGCTCCGGGAGATCCAGATCTCTGCTGCGCTTCCGCC 306
|||||

QY 301 ATGAGAAATGTCATCGGCATCCGAGACATTTCTCGGGGCGTCCACCTTGGAGCCATGAGAG 360
|||||
Db 307 ATGAGAAATGTCATCGGCATCCGAGACATTTCTCGGGGCGTCCACCTTGGAGCCATGAGAG 366
|||||

QY 361 ATGTCTACATTTGTCGAGGACCTGTAGGAGACTGACTGTACAAAGTTGCTGAAAAAGCCAGC 420
|||||
Db 367 ATGTCTACATTTGTCGAGGACCTGTAGGAGACTGACTGTACAAAGTTGCTGAAAAAGCCAGC 426
|||||

QY 421 AGCTGAGCAATGACCATATCTCTACTTCTCTTACCAGATCTCTGCGGGGCTCAAGTACA 480
|||||
Db 427 AGCTGAGCAATGACCATATCTCTACTTCTCTTACCAGATCTCTGCGGGGCTCAAGTACA 486
|||||

QY 481 TCCACTCCGCCAAACGTGCTCCACCGAGATCTTAAAGCCCTTCAACCTGCTCATCAACACCA 540
|||||
Db 487 TCCACTCCGCCAAACGTGCTCCACCGAGATCTTAAAGCCCTTCAACCTGCTCATCAACACCA 546
|||||

QY 541 CTTGGGACCTTAAGATTTTGTGATTTTCGGCTCGGCCCGGATTCGCCATCTGAGCATGACC 600
|||||
Db 547 CTTGGGACCTTAAGATTTTGTGATTTTCGGCTCGGCCCGGATTCGCCATCTGAGCATGACC 606
|||||

QY 601 ACACCGGCTTCTCGACGGAGTATGTGGCTACGCGCTGCTGTTACCGGGCCCCAGAGATCATGC 660
|||||
Db 607 ACACCGGCTTCTCGACGGAGTATGTGGCTACGCGCTGCTGTTACCGGGCCCCAGAGATCATGC 666
|||||

QY 661 TGAATCTCAAGGGCTATACCAAGTCCATGACATCTGTGCTGTGGCTGCATTTGGGCTG 720
|||||
Db 667 TGAATCTCAAGGGCTATACCAAGTCCATGACATCTGTGCTGTGGCTGCATTTGGGCTG 726
|||||

QY 721 AGATGCTCTTAACCGGCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACCA 780
|||||
Db 727 AGATGCTCTTAACCGGCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACCA 786
|||||

QY 781 TTCTCGGATCTCTGGGCTCCCATCCAGGAGACCTGAATTTGATCATCAACATGAAGG 840
|||||
Db 787 TTCTCGGATCTCTGGGCTCCCATCCAGGAGACCTGAATTTGATCATCAACATGAAGG 846
|||||

QY 841 CCGGAAATACCTACAGTCTCTGCCCTCCAAAGCAAGTGGCTTGGGCAAGCTTTTTC 900
|||||
Db 847 CCGGAAATACCTACAGTCTCTGCCCTCCAAAGCAAGTGGCTTGGGCAAGCTTTTTC 906
|||||

QY 901 CCAAGTCAGATCCAAAGCCCTTGACCTCTCGACCGGATGTTAACTTTAAACCCCAATA 960
|||||
Db 907 CCAAGTCAGATCCAAAGCCCTTGACCTCTCGACCGGATGTTAACTTTAAACCCCAATA 964
|||||

QY 961 AACGGATCACAGTGGAGGA 979
|||||
Db 965 AAMGGGTACAGTGGRGA 983
|||||
```

```
RESULT 15
BX445445/c 1108 bp mRNA linear EST 03-MAY-2004
LOCUS BX445445 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CS0DN005YAI4 3-PRIME, mRNA sequence.
ACCESSION BX445445
VERSION BX445445.2 GI:46958964
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1108)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30770386.
Contact: Genoscope
Genoscope - Centre National de Sequencage
```


2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 8342.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1AN002ZA08NP1&c=8342.f.

FEATURES
Location/Qualifiers
1. .1108
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN005YA14"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN		Query Match	Score	DB 5;	Length	1108;
		Best Local Similarity	93.48;	Pred. No. 8.7e-204;		
		Matches 1065;	Conservative 3;	Mismatches 1;	Indels 71;	Gaps 5;
Qy	691	ACATCTGGTCTGGGCTGCAATTCGGCTGAGATGCTCTTAACGGGCCCATCTTCCCTG	750			
Db	1087	ACATCTGGTCTGGGCTGCAATTCGGCTGAGATGCTCTTAACGGGCCCATCTTCCCTG	1029			
Qy	751	GCAGCACTACCTGGATGCTCAACACATTCCTGGGCTCCCTGGCTCCCATCCAGG	810			
Db	1028	GCAAGCACTACCTGGATGCTCAACACATTCCTGGGCTCCCTGGCTCCCATCCAGG	969			
Qy	811	AGGACCTGAAATTTGATCATCAACATGAAGGCCCGA-AACTACCTACAGTCTCTGCCCTCC	869			
Db	968	AGGACCTGAAATTTGATCATCAACATGAAGGCCCGA-AACTACCTACAGTCTCTGCCCTCC	909			
Qy	870	AAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGACTCCAA	917			
Db	908	AAGACCAAGGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGACTCCAAAGTTTCGAGAGAC	849			
Qy	918	-----GCCCTTGACTGCTGACCGGATGTTAACTTTAACCCCAATAAAGGATCACAG	972			
Db	848	ATGAGGCCCTTGACTGCTGACCGGATGTTAACTTTAACCCCAATAAAGGATCACAG	789			
Qy	973	TGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACGGATGAGCCAG	1032			
Db	788	TGGAGGAAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGACGGATGAGCCAG	729			
Qy	1033	TGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTAAGAGGCGGTGA	1092			
Db	728	TGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTAAGAGGCGGTGA	669			
Qy	1093	AGGAGCTCATCTTCAGGAGACAGCAGCTTCAGCCCGGAGTGTGGAGGCCCTTAGC	1152			
Db	668	AGGAGCTCATCTTCAGGAGACAGCAGCTTCAGCCCGGAGTGTGGAGGCCCTTAGC	609			
Qy	1153	CCAGACAGACATCTCTGACCCCTGGGCTCGAACAGAACTGGCAAGAGGCAAGAGTC	1212			
Db	608	CCAGACAGACATCTCTGACCCCTGGGCTCGAACAGAACTGGCAAGAGGCAAGAGTC	579			
Qy	1213	ACTGAGGCGCTCTGTACCCAGGACCTTGCCTCTCTGCTGCCCTCTCCGCCAGACTGTT	1272			
Db	578	-----GGACCTGCTCTCTGCTGCCCTCTCCGCCAGACTGTT	540			
Qy	1273	AGAAATGGACATCTGTGCCCAGCCCGGACCTTGGCAGCCAGGCGGGGTGGAGCATGGG	1332			
Db	539	AGAAATGGACATCTGTGCCCAGCCCGGACCTTGGCAGCCAGGCGGGGTGGAGCATGGG	480			

Search completed: February 7, 2006, 14:53:00
Job time : 7241.32 secs

Qy	1333	CCTGGCCACCTCTCTCTCTTGTGCTGAGGCTTCAGGCTTCAGGAGGCAAGGCCCTTCTCCT	1392
Db	479	CCTGGCCACCTCTCTCTCTTGTGCTGAGGCTTCAGGCTTCAGGAGGCAAGGCCCTTCTCCT	420
Qy	1393	CCCCACCCGCGCTCCCAAGCGGCGCTCGGAGCTCAGGTGGCCCCCAGTTCAATCTCCCGC	1452
Db	419	CCCCACCCGCGCTCCCAAGCGGCGCTCGGAGCTCAGGTGGCCCCCAGTTCAATCTCCCGC	360
Qy	1453	TGCTGTGCTGGCGCCCTTACCTTCCCCAGCGTCCCAAGTCTCTGGCAGTCTTGGAAATGAA	1512
Db	359	TGCTGTGCTG-TGGCGCCCTTACCTTCCCCAGCGTCCCAAGTCTCTGGCAGTCTTGGAAATGAA	301
Qy	1513	GGTTTCTGGCTGGCCCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGGGGGCGCTGAGTAG	1572
Db	300	GGTTTCTGGCTGGCCCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGGGGGCGCTGAGTAG	241
Qy	1573	GGAATCAGGGCCATGCTGCTCCCGCCCTCATCTCATTCAAACCCCAACCTAGTTTCCCTGAA	1632
Db	240	GGAATCAGGGCCATGCTGCTCCCGCCCTCATCTCATTCAAACCCCAACCTAGTTTCCCTGAA	181
Qy	1633	GGAACATTCCTTAGTCTCAAGGGCTAGCATTCCTGAGGAGCCAGGCGGGCGGAGTCCCC	1692
Db	180	GGAACATTCCTTAGTCTCAAGGGCTAGCATTCCTGAGGAGCCAGGCGGGCGGAGTCCCC	121
Qy	1693	TCCTGTCAAAGCTGTCACTTCGCGTGGCCCTCGCTGCTTCTGTGTGAGCAGAGT	1752
Db	120	TCCTGTCAAAGCTGTCACTTCGCGTGGCCCTCGCTGCTTCTGTGTGAGCAGAGT	61
Qy	1753	GGAAGCTGGGGGGCGTGGAGAGCCCGCGCCCTTGCCACCTTCCCTGACCCCTCTAATATAT	1812
Db	60	GGAAGCTGGGGGGCGTGGAGAGCCCGCGCCCTTGCCACCTTCCCTGACCCCTCTAATATAT	1

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:30:54 ; Search time 391.141 Seconds
(without alignments)
8348.349 Million cell updates/sec

Title: US-10-623-108-5

Perfect score: 1837

Sequence: 1 gagagtgagatggcgccg.....tagagatgtctatggctg 1837

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCPUS_COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1708.4	93.0	1866	3	US-08-909-742-1
2	1708.4	93.0	1866	3	US-09-412-289-1
3	1138.4	62.0	1896	3	US-08-417-197-38
4	1025	55.8	1747	2	US-08-176-620A-1
5	1025	55.8	1747	2	US-08-463-862-1
6	1025	55.8	1747	2	US-08-461-985-1
7	1025	55.8	1747	2	US-08-458-887-1
8	1025	55.8	1747	3	US-08-932-787B-1
9	1025	55.8	1747	3	US-08-932-012C-1
10	1025	55.8	1747	3	US-08-888-818C-1
11	958.6	32.5	1875	3	US-09-417-197-56
12	651.4	35.5	1701	3	US-09-949-016-4048
13	646.6	35.2	1611	3	US-08-909-742-2
14	646.6	35.2	1611	3	US-09-412-289-2
15	642	34.9	2791	3	US-09-016-434-1263
16	632.4	34.4	1467	2	US-08-176-620A-3
17	632.4	34.4	1467	2	US-08-463-862-3
18	632.4	34.4	1467	2	US-08-461-985-3
19	632.4	34.4	1467	2	US-08-458-887-3
20	632.4	34.4	1467	3	US-08-932-787B-3
21	632.4	34.4	1467	3	US-08-932-012C-3
22	632.4	34.4	1467	3	US-08-888-818C-3
23	632.4	34.4	1815	3	US-09-417-197-58
24	632.4	34.4	1818	3	US-09-417-197-40

c	25	525.8	28.6	546	3	US-09-854-133-457	Sequence 457, App
	26	494	26.9	574	3	US-09-854-133-456	Sequence 456, App
	27	454.6	24.7	549	2	US-08-176-620A-7	Sequence 7, Appli
	28	454.6	24.7	549	2	US-08-461-985-7	Sequence 7, Appli
	29	454.6	24.7	549	3	US-08-932-787B-15	Sequence 15, Appl
	30	454.6	24.7	549	3	US-08-932-012C-15	Sequence 15, Appl
	31	454.6	24.7	549	3	US-08-888-818C-15	Sequence 15, Appl
	32	383.4	20.9	537	2	US-08-176-620A-9	Sequence 9, Appli
	33	383.4	20.9	537	2	US-08-461-985-9	Sequence 9, Appli
	34	383.4	20.9	537	3	US-08-932-787B-17	Sequence 17, Appl
	35	383.4	20.9	537	3	US-08-932-012C-17	Sequence 17, Appl
	36	383.4	20.9	537	3	US-08-888-818C-17	Sequence 3926, Ap
	37	320.4	17.4	2826	3	US-09-949-016-3926	Sequence 1458, Ap
	38	320.4	17.4	2828	3	US-09-016-434-1458	Sequence 11, Appl
	39	320.4	17.4	2980	3	US-09-266-225D-11	Sequence 1, Appli
	40	312.4	17.0	1502	2	US-08-651-940-1	Sequence 1, Appli
	41	312.4	17.0	1502	3	US-09-295-029-1	Sequence 1, Appli
	42	312.4	17.0	1502	3	US-09-724-768-1	Sequence 1, Appli
	43	307.6	16.7	1334	3	US-09-949-016-3997	Sequence 3997, Ap
	44	307.6	16.7	1310	3	US-09-047-288-1	Sequence 1, Appli
	45	307.6	16.7	1310	3	US-08-802-191-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-909-742-1

; Sequence 1, Application US/08909742

; Patent No. 6007991

; GENERAL INFORMATION:

; APPLICANT: Vimala S. Sivaraman

; APPLICANT: Hsien-Yu Wang

; APPLICANT: Craig C. Malbon

; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES FOR MITOGEN-

; TITLE OF INVENTION: ACTIVATED PROTEIN KINASES AS THERAPY FOR

; TITLE OF INVENTION: BREAST CANCER

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann & Barton, LLP

; STREET: 350 Jericho Turnpike

; CITY: Jericho

; STATE: New York

; COUNTRY: USA

; ZIP: 11753

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for windows

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/909,742

; FILING DATE: August 12, 1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/831,994

; FILING DATE: April 1, 1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/827,520

; FILING DATE: March 28, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Adams, Lindsay S.

; REGISTRATION NUMBER: 36,425

; REFERENCE/DOCKET NUMBER: 178-225 CIP II

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 822-3550

; TELEFAX: (516) 822-3582

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1866 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA									
; HYPOTHETICAL: NO									
; ANTI-SENSE: NO									
US-08-909-742-1									
Query Match 93.0%; Score 1708.4; DB 3; Length 1866;									
Best Local Similarity 97.0%; Pred. No. 0;									
Matches 1783; Conservative 0; Mismatches 1; Indels 54; Gaps 2;									
QY	3	GGAGTGGAGATGCGCGCGCGCGGCTCAGGGGGGCGGGGGCGGGGAGCCCGGTAGAAC	62						
DB	64	GGAGTGGAGATGCGCGCGCGCGGCTCAGGGGGGCGGGGGCGGGGAGCCCGGTAGAAC	123						
QY	63	GAGGGGTCGGCCCGGGGGTCCCGGGGAGTGGAGATGGTGAAGGGGAGCGGTTTCGAC	122						
DB	124	GAGGGGTCGGCCCGGGGGTCCCGGGGAGTGGAGATGGTGAAGGGGAGCGGTTTCGAC	183						
QY	123	GTGGGCGCGGCTACACGCACTGTCAGTATCATCGGCGAGGGCGGTACCGGCATGTGTCAG	182						
DB	184	GTGGGCGCGGCTACACGCACTGTCAGTATCATCGGCGAGGGCGGTACGGCATGTGTCAG	243						
QY	183	TCGGGCTATGACCACTGTGGCGAAGACTCGGCTGGCCATCAAGAGATCAGCCCTTCGAA	242						
DB	244	TCGGGCTATGACCACTGTGGCGAAGACTCGGCTGGCCATCAAGAGATCAGCCCTTCGAA	303						
QY	243	CATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCTGCTGGCTTCGGCCAT	302						
DB	304	CATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCTGCTGGCTTCGGCCAT	363						
QY	303	GAGATGTTCATCGGCGATCCGAGACATTCCTGCGGGCGTCCACCTCGGAAGCCATGAGAGAT	362						
DB	364	GAGATGTTCATCGGCGATCCGAGACATTCCTGCGGGCGTCCACCTCGGAAGCCATGAGAGAT	423						
QY	363	GTCTCATTTGTGAGGACCTGATGGAGACTGACCTGTACAGTTGCTGAAAGCCAGAG	422						
DB	424	GTCTCATTTGTGAGGACCTGATGGAGACTGACCTGTACAGTTGCTGAAAGCCAGAG	483						
QY	423	CTGAGCAATGACCATCTGCTACTTCTCTACAGATCTCGGGGCGCTCAAGTACATC	482						
DB	484	CTGAGCAATGACCATCTGCTACTTCTCTACAGATCTCGGGGCGCTCAAGTACATC	543						
QY	483	CACCTCGCCCAACGTGTCTCAACGAGATCTAAAGCCCTCCAACTGCTCATCAACACACC	542						
DB	544	CACCTCGCCCAACGTGTCTCAACGAGATCTAAAGCCCTCCAACTGCTCAGCAACACACC	603						
QY	543	TGGGACCTTAAGATTTGTGATTTGGGCTTGGCCCGGATTCGCGATCCTGAGCATGACAC	602						
DB	604	TGGGACCTTAAGATTTGTGATTTGGGCTTGGCCCGGATTCGCGATCCTGAGCATGACAC	663						
QY	603	ACCGGCTTCTGACGAGATATGGGCTACGCGTGTGACGGGCCCGGAGATCATGCTG	662						
DB	664	ACCGGCTTCTGACGAGATATGGGCTACGCGTGTGACGGGCCCGGAGATCATGCTG	723						
QY	663	AACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCAATCTGGCTGAG	722						
DB	724	AACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCAATCTGGCTGAG	783						
QY	723	ATGCTCTCTAAACCGGGCCATCTTCCCTGGCAAGCACTACCTGATCAGCTCAACCAATT	782						
DB	784	ATGCTCTCTAAACCGGGCCATCTTCCCTGGCAAGCACTACCTGATCAGCTCAACCAATT	843						
QY	783	CTGGGCTATCTGGGCTCCCATCCAGGAGGACCTGAATTTGTATCATCAACATGAAGGCC	842						
DB	844	CTGGGCTATCTGGGCTCCCATCCAGGAGGACCTGAATTTGTATCATCAACATGAAGGCC	903						
QY	843	CGAAACTACCTACAGTCTCTGCGCTCCAAAGCAAGGTGGCTTGGGCCAAGCTTTTCCCC	902						
DB	904	CGAAACTACCTACAGTCTCTGCGCTCCAAAGCAAGGTGGCTTGGGCCAAGCTTTTCCCC	963						
QY	903	AAGTCAGCTCAAGCCCTTGACCTGTGGACCGGATGTAACCTTTAACCCCAATAA	962						
DB	964	AAGTCAGCTCAAGCCCTTGACCTGTGGACCGGATGTAACCTTTAACCCCAATAA	1023						

QY	963	CGGATCACAGTGGAGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGAG	1022
DB	1024	CGGATCACAGTGGAGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGAG	1083
QY	1023	GATGAGCCAGTGGCCGAGAGCCCTTACCTTTCGATGGAGCTGGAGCTGATCTACTAAG	1082
DB	1084	GATGAGCCAGTGGCCGAGAGCCCTTACCTTTCGATGGAGCTGGAGCTGATCTACTAAG	1143
QY	1083	GAGCGGCTGAGGAGCTCATCTTCCAGGAGCAGACGCTTCCAGCCGAGCTGCTGGAG	1142
DB	1144	GAGCGGCTGAGGAGCTCATCTTCCAGGAGCAGACGCTTCCAGCCGAGCTGCTGGAG	1203
QY	1143	GGCCCCCTAGCCCAAGACATCTCTGACCCCTGGGGGCTCGAAACAGAACTGGCAAGAG	1202
DB	1204	GGCCCCCTAGCCCAAGACATCTCTGACCCCTGGGGGCT-----	1243
QY	1203	GCAAGAGTCACTGAGGGCTCTGTACCCAGGACCTGCTCTCTGCTGCCCCCTCTCCG	1262
DB	1244	-----GGACCTGCTCTGCTGCCCCCTCTCCG	1272
QY	1263	CCAGACTGTTAGAAATGGACACTGTGCCAGCCGCGACCTTGGCAGCCAGCCGCGGT	1322
DB	1273	CCAGACTGTTAGAAATGGACACTGTGCCAGCCGCGACCTTGGCAGCCAGCCGCGGT	1332
QY	1323	GGAGCATGGGCTGGCCACCTCTCTCTTGTGTAGGCTCCAGCTTCAGGCAGGCCAAG	1382
DB	1333	GGAGCATGGGCTGGCCACCTCTCTCTTGTGTAGGCTCCAGCTTCAGGCAGGCCAAG	1392
QY	1383	GCCTTCTCTCCCAACCGCCCTCCCAAGGGGCTCGGGAGCTCAGTGGGCCCAAGTTC	1442
DB	1393	GCCTTCTCTCCCAACCGCCCTCCCAAGGGGCTCGGGAGCTCAGTGGGCCCAAGTTC	1452
QY	1443	AACTCTCCC---GCTGCTGCTGCTGCGCCCTTACCTTCCCAGGCTCCAGTCTCTGGCAG	1499
DB	1453	AACTCTCCC---GCTGCTGCTGCTGCGCCCTTACCTTCCCAGGCTCCAGTCTCTGGCAG	1512
QY	1500	TTCTGGAATGGAAGGTTCTGGCTGCCCCCAACCTGCTGAAGGGCAGAGTGGAGGTGG	1559
DB	1513	TTCTGGAATGGAAGGTTCTGGCTGCCCCCAACCTGCTGAAGGGCAGAGTGGAGGTGG	1572
QY	1560	GGGCGTGAATGGAAGTCTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAAACCCACCC	1619
DB	1573	GGGCGTGAATGGAAGTCTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAAACCCACCC	1632
QY	1620	TAGTTTCCCTGAAGCAATTCCTTAGTCTCAAGGCTAGCATCCTGAGGAGCCAGGCC	1679
DB	1633	TAGTTTCCCTGAAGCAATTCCTTAGTCTCAAGGCTAGCATCCTGAGGAGCCAGGCC	1692
QY	1680	GGGCGGAATCCCCTCTGTCAAAGCTGTCACTTCGCGTCCCTCGCTGCTTCTGTGT	1739
DB	1693	GGGCGGAATCCCCTCTGTCAAAGCTGTCACTTCGCGTCCCTCGCTGCTTCTGTGT	1752
QY	1740	GGTGAGCAAGTGGAGCTGGGGGGCTGGAGAGCCCGGGCCCTGCCACCTCCCTGAC	1799
DB	1753	GGTGAGCAAGTGGAGCTGGGGGGCTGGAGAGCCCGGGCCCTGCCACCTCCCTGAC	1812
QY	1800	CGGCTCTAATATAAATATAGAGATGTGTCTATGGCTG	1837
DB	1813	CGGCTCTAATATAAATATAGAGATGTGTCTATGGCTG	1850

RESULT 2
US-09-412-289-1
; Sequence 1, Application US/09412289
; Patent No. 6271210
; GENERAL INFORMATION:
; APPLICANT: Sivaraman, Vimala S.
; APPLICANT: Wang, Hsien-Yu
; APPLICANT: Malbon, Craig C.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES FOR MITOGEN-ACTIVATED
; PROTEIN KINASES AS THERAPY FOR CANCER
; FILE REFERENCE: Seq. 1-4 (178-225 CIP 11/CON)
; CURRENT APPLICATION NUMBER: US/09/412,289

RESULT 3									
US-09-417-197-38									
; Sequence 38, Application US/09417197									
; Patent No. 6518021									
; GENERAL INFORMATION:									
; APPLICANT: Ole THASTRUP, et al.									
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I									
; FILE REFERENCE: 3759-0110P									
; CURRENT APPLICATION NUMBER: US/09/417,197									
; NUMBER OF FILING DATE: 1999-10-07									
; NUMBER OF SEQ ID NOS: 143									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 38									
; LENGTH: 1896									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: EGFP-Erk1 fusion construct									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (1)..(1893)									
US-09-417-197-38									
Query Match									
Best Local Similarity 62.0%; Score 1138.4; DB 3; Length 1896;									
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	12	ATGCGCGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCGTAGAACCCGAGGGGGTTC	71						
DB	757	ATGCGCGCGCGCGCGCTCAGGGGGCGGGGGCGGGAGCCCGTAGAACCCGAGGGGGTTC	816						
QY	72	GGCCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCGCTTCAGCTGGGCGCG	131						
DB	817	GGCCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCGCTTCAGCTGGGCGCG	876						
QY	132	CGCTACACGAGTTCGAGTACATCGCGGAGGGCGGTACGGATGGTTCAGCTCGGGCTAT	191						
DB	877	CGCTACACGAGTTCGAGTACATCGCGGAGGGCGGTACGGATGGTTCAGCTCGGGCTAT	936						
QY	192	GACACGTCGGCAAGACTCGCTGGGCCATCAAGAAGATCAGGCCCTTCGAAACATCAGACC	251						
DB	937	GACACGTCGGCAAGACTCGCTGGGCCATCAAGAAGATCAGGCCCTTCGAAACATCAGACC	996						
QY	252	TACTGCGGCGAGCTTCGGGAGATCCAGATCCTGCTGGCTTCGCCCATGAGATGTC	311						
DB	997	TACTGCGGCGAGCTTCGGGAGATCCAGATCCTGCTGGCTTCGCCCATGAGATGTC	1056						
QY	312	ATCGGCATCCGAGACATTCCTGGGGGCTCCACCCTCGAAGCCATGAGATGTCTACATT	371						
DB	1057	ATCGGCATCCGAGACATTCCTGGGGGCTCCACCCTCGAAGCCATGAGATGTCTACATT	1116						
QY	372	GTGAGGACCTGATGGAGACTGACCTGTACAGTTGCTGAAAAGCCAGCAGCTGAGCAAT	431						
DB	1117	GTGAGGACCTGATGGAGACTGACCTGTACAGTTGCTGAAAAGCCAGCAGCTGAGCAAT	1176						
QY	432	GACCATATCTGCTACTTCTTACAGATCTCTGGGGGCTCAAGTACATCACTCCGCG	491						
DB	1177	GACCATATCTGCTACTTCTTACAGATCTCTGGGGGCTCAAGTACATCACTCCGCG	1236						
QY	492	AACGTGCTCCAGGAGATCTAAGCCCTTCAACCTGCTCATCAACACCACTGCGACCTT	551						
DB	1237	AACGTGCTCCAGGAGATCTAAGCCCTTCAACCTGCTCATCAACACCACTGCGACCTT	1296						
QY	552	AAGATTGTGATTTCGGCCCTGGCCCGGATTTGCCGATCCTGAGCATGACACACCGGCTTC	611						
DB	1297	AAGATTGTGATTTCGGCCCTGGCCCGGATTTGCCGATCCTGAGCATGACACACCGGCTTC	1356						
QY	612	CTGACCGAGTATGTGCTACGCGCTGGTACCGGGCCCGAGATCATGTGACTCCAG	671						
DB	1357	CTGACCGAGTATGTGCTACGCGCTGGTACCGGGCCCGAGATCATGTGACTCCAG	1416						
QY	672	GGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCAATTCGGCTGAGATGCTCTCT	731						

RESULT 4

US-08-176-620A-1

; Sequence 1, Application US/08176620A

; Patent No. 5595904

GENERAL INFORMATION:

APPLICANT: Boulton, Teri G.

APPLICANT: Cobb, Melanie H.

APPLICANT: Yancopoulos, George D.

APPLICANT: Nye, Steven

APPLICANT: Panayotatos, Nikos

TITLE OF INVENTION: A Family of Map2 Protein Kinases

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/176,620A

FILING DATE: 03-JAN-1994

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Misrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 6526-123

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1747 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

```

; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1095
; US-08-176-620A-1

Query Match      55.8%; Score 1025; DB 2; Length 1747;
Best Local Similarity 78.1%; Pred. No. 2.1e-202;
Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

QY 58 GAACCGAGGGGTCGGCCCGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGACGGCT 117
DB 5 GAACTGCTGGGGTGGTCCCGGTGGTCCCGGGGAGGTGGAGGTGGTGAAGGGGACGGCAT 64

QY 118 TCAGAGTGGGCGCGCTACACGAGTTCAGTACATCGCGGAGGGCGGTACGGATGG 177
DB 65 TCAGAGTGGGCGCGCTACACGAGTTCAGTACATCGCGGAGGGCGGTACGGATGG 124

QY 178 TCAGCTCGGCTATGACCACTGCGCAAGACTCGCGTGGCCATCAAGAAGATCAGCCCT 237
DB 125 TCAGCTCAGCATATGACCACTGCGCAAGACTCGAGTGGCTATCAAGAAGATCAGCCCT 184

QY 238 TCGAATCATAGACTTACGCGGAGCGCTCGCGGAGATCCAGATCCTGCTGGCTTC 297
DB 185 TCAGCATCAAACTACTGTGAGCGCACTGCTGAGAGAAATCCAGATCTTCTCGATTC 244

QY 298 GCCATGAGATGTCATCGGATCCGAGACATTCGCGGGGCTCCACCCTCGAAGCCATGA 357
DB 245 GCCATGAGATGTCATAGGATCCGAGACATTCCTCAGAGACCCACCCTCGAAGCCATGA 304

QY 358 GAGATGTCATATGTCGAGGACCTGATGGAGACTGACCTGTGTACAACTTCTGAAAGCC 417
DB 305 GAGATGTCATATGTCGAGGACCTGATGGAGACTGACCTGTGTACAACTTCTGAAAGCC 364

QY 418 AGCAGCTGAGCAATGACCAATATCTGCTATCTTCTTACAGATCTCGCGGGGCTCAAGT 477
DB 365 AGCAGCTGAGCAATGACCAATATCTGCTATCTTCTTACAGATCTCGCGGGGCTCAAGT 424

QY 478 ACATCACTCGGCAAGCTGCTCAGCAGATCTAAGGACCTCCAACTGCTCATCAACA 537
DB 425 ACATCACTCGGCAAGCTGCTCAGCAGATCTAAGGACCTCCAACTGCTCATCAACA 484

QY 538 CCACCTCGACCTTAAGATTTGATTTGCGGCTCGGCGGATTCGCGATCTGAGCATG 597
DB 485 CCACCTCGACCTTAAGATCTGATTTGCGGCTCGGCGGATTCGCGATCTGAGCATG 544

QY 598 ACCACACCGGCTTCTGACCGAGTATGCGCTACGCGTGTGTACCGGGGCGGAGATCA 657
DB 545 ACCACACCGGCTTCTGACCGAGTATGCGCTACGCGTGTGTACCGGGGCGGAGATCA 604

QY 658 TGCTGAATCCAGGGCTATACCAAGTCCATCGACATCTGCTGCTGGGCTGCAATCTGG 717
DB 605 TGCTGAATCCAGGGCTATACCAAGTCCATCGACATCTGCTGCTGGGCTGCAATCTGG 664

QY 718 CTGAGATGCTCTTAACCGGCGCTATCTTCCCTGCGAGCACTACTCGATCAGCTCAACC 777
DB 665 CTGAGATGCTCTTAACCGGCGCTATCTTCCCTGCGAGCACTACTCGATCAGCTCAACC 724

QY 778 ACATTTCTGGGCTATCTGGGCTCCCACTCCAGGAGGACCTGAATTTGATATCAACATGA 837
DB 725 ACATTTCTGGGCTATCTGGGCTCCCACTCCAGGAGGACCTGAATTTGATATCAACATGA 784

QY 838 AGGCGCGAACTACTACAGTCTCTGCGCTCCAGACCAAGTGGCTGGGCCAAGCTTT 897
DB 785 AGGCGCGAACTACTACAGTCTCTGCGCTCCAGACCAAGTGGCTGGGCCAAGCTTT 844

QY 898 TCCCAAGTCTGAGCTCCAAAGCCCTTGAACCTGCTGGACCGATGTAAACCTTTAACCCCA 957
DB 845 TCCCAAGTCTGAGCTCCAAAGCCCTTGAACCTGCTGGACCGATGTAAACCTTTAACCCCA 904

QY 958 ATAAACGGATCAGTGGAGGAGCGGTGGCTCACCCTTACCTGGAGCAGTACTATGACC 1017
DB 1017 ATAAACGGATCAGTGGAGGAGCGGTGGCTCACCCTTACCTGGAGCAGTACTATGACC

```

RESULT 5

```

US-08-463-862-1
; Sequence 1, Application US/08463862
; Patent No. 5776751
;
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G. et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

```

```

DB 905 ACAGCGCATCACAGTAGAGGAAGCACTGGCTCACCCCTTACCTTGGAAACAGTACTATGATC 964
QY 1018 CGACGGATCAGCGAGTGGCGGAGAGCCCTTACCTTCCGCATGGAGCTGGATGACCTAC 1077
DB 965 CGACAGATGAACCAAGTGGCTGAGGAGCCATTCACCTTTGACATGGAGCTGGATGATCCTCC 1024
QY 1078 CTAAGGAGCGGCTGAAGGAGCTCATCTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGC 1137
DB 1025 CCAAGGAGCGGCTGAAGGAGCTGATCTTCAAGAGACAGCCCTTCCAGCCAGGGGCAC 1084
QY 1138 TGAGGCGCCCTAGCCCGACAGACATCTCTGACCCCTGGGGCTTGAAACAGAACTGGCA 1197
DB 1085 CAGAGGCCCTTAAACAAAGAACACACACCTCTGCTTTTGGACCTGG----- 1131
QY 1198 AAGAGCAAGAGGTCACTGAGGCGCTCTGTACCCAGGACCTGCTCTGCTGCTGCTGCTCCT 1257
DB 1132 -----TCTGCTCTACCTGCTCCT 1149
QY 1258 TCCCGCCAGACTGTTAGAAATGGACACATGTGCCCCAGCCCGGACCTTGGCAGCCCGAGGCC 1317
DB 1150 TCTCTCGAGATGTTAGAAATG--AACTTTGTCTCAACCCGAGACCCCGGAGCCCGAGG--C 1206
QY 1318 GGGGTGGAGCATGGGCTGCGCCACCTCTCTCTTTGCTGAGGCTTCCAGTTCAGGCAGG 1377
DB 1207 TGAACCAAGGTGGGCTGGCACCTCTCTCACTCTGCTGGGCTCTCTCTGTTCAAG--AG 1264
QY 1378 CCAAGGCTTCTCTCCCAACCGCCCTCCCAAGGGGCTCGGAGCTCAGTGGGCCCC 1437
DB 1265 GCTTCTCCACCTCAGTCCCTGCCCCATCTCCCTTTGACCTGAGTGAAGGTGGTCCC 1324
QY 1438 AGTTCAATCTCCGCTGCTGCTGCTGCGGCTTACCTTCCCCAGCGTCCCAGTCTCTGGC 1497
DB 1325 AGAGCTGATCTCTGCTGCTG--TGTCTTTATCTATCTGCTAGCCCCAGCTCTGGTAGAC 1383
QY 1498 AGTTCTGGAATGAAGGTTCTGGCTGCCCCAACCTGTGAAAGGCGAGAGTGGAGGGTG 1557
DB 1384 GGTCTGGAATGAAGGCTATGACCGCCCTA-----GACCTGTGCTACAGAGGGTG 1437
QY 1558 GGGGGGCTGAGTAGGAGCTCAGGGGCTGCGGCTGCGGCTCATCTCATTTCAAACCCAC 1617
DB 1438 GAGGCGACTGAGT-----AGGCTAAGCTCTGCGCTACTCATCTGTTGGAACCCAC 1489
QY 1618 CTTAGTTTCTGAAAGAACTTCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGG 1677
DB 1490 CCAATTTTCTGACAGAACTTCTTAAATCTCAAGGCTAGTTTCCCTGAGGAGCCAGC 1549
QY 1678 CCGGGCGGATCCCTCCCTGTCAAAGCTGTCACTTCCGCTGCGCTGCTGCTTCTGT 1737
DB 1550 CTAGGCTTAACTCTCCCTCTC--AAGCTGCCATGTAACGCTTGTGCTTCTGT 1608
QY 1738 GT--GGTGAAGAGTGGAGCTGGGGGCGTGGAGAGCCCGGCGGCTGCGACCTCCCT 1796
DB 1609 GTGGTGAATGATGTTGAGGCGGGGCGCTGGAGAGCCCGCTGCGGCTGCGGCTCCCTCCCT 1668
QY 1797 GAGCC---GTCTAATATATAATATAGAGATGTCTATGG 1834
DB 1669 GTGCTGTATCTAATATATAATATAGAGATGTGTATATGG 1709

```

```

; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,862
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,544
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1095
; US-08-463-862-1

```

Query Match	55.8%;	Score 1025;	DB 2;	Length 1747;
Best Local Similarity	78.1%;	Pred. No. 2.1e-202;		
Matches 1391;	Conservative	0;	Mismatches 310;	Indels 80;
				Gaps 10;

58	QY	GAACCGAGGGGGTCCGCCCGGGGGTCCGGGGAGGTGGAGATGGTGAAGGGCGAGCGGT	117
5	DB	GAATCTGTCGGGGTCCGTCCCGGTGGTCCCGGGGAGGTGGAGGTGGTGAAGGGCGAGCGAT	64
118	QY	TCGACGTGGGCCCGCGCTTACACGCAGTTGTCAGTATCATCGGCGAGGGGGCGCTACGGCATGG	177
65	DB	TCGACGTGGGCCCAACGCTACACGCAGCTGTCAGTATCATCGGCGAGGGCGCGTACGGCATGG	124
178	QY	TCAGCTCGGCCCTATGACCAACGTGGCGCAAGTATCGCGTGGCCATCAAGAAGATCAGCCCT	237
125	DB	TCAGCTCAGCATATGACACACGTGCGCAAGACACAGATGGCTATCAAGAAGATCAGCCCT	184
238	QY	TCGAAACATCAGACCTATCTGCAGCGCAGCTTCGGGAGATCCAGATCTCTGCTGGCTTCC	297
185	DB	TCGAGCATCAAACTACTGTTCAGCGCAGCTTGAGAGAAATCCAGATCTTGTTCGGATTCC	244
298	QY	GCCATGAGAATGTCATCGGCATCCGAGACATCTTCGGGGCGTCCACCTCGAAGCCATGA	357
245	DB	GCCATGAGAATGTCATAGGCATCCGAGACATCTTCAGAGCACCCACCTTCGAAGCCATGA	304
358	QY	GAGATGTCATATTTGTGCAGACCTGATGGAGACTGACCTGTCAAGATTGCTGAAAAGCC	417
305	DB	GAGATGTTTACATTTGTTTCAGAACCTCATGGAGACGGACCTGTCAAGCTGTCTAAAGACC	364
418	QY	AGCAGCTGAGCAATGACCAATATCTGCTACTTTCCTCTACAGATCTCTGGGGGCTCTAAGT	477
365	DB	AGCAGCTGAGCAATGACCAATCTGCTACTTTCCTCTACAGATCTCTGGGGGCTCTAAGT	424
478	QY	ACATCCACTCGCCCAAGTGTCTCCACCGAGATCTTAAGCCCTCCAACTGCTCATCAACA	537
425	DB	ACATACACTCGGCCCAATGTGCTGACCGGGACCTGAAGCCCTTCCAATCTGCTTTATCAACA	484
538	QY	CCACTCGGACCTTTAAGATTGTGATTTTCGGCCCTGGCCCGGATTTGCCGATCTCTGAGCATG	597
485	DB	CCACTCGGACCTTTAAGATCTGTGATTTTGGCCCTTGGCCCGATTTGCTGACCTTGAGCAG	544


```

Db      1550 CTAGGCGCTAACCCCTCTCCCTCTC-AAAGTGCACATGTAAAGCCCTTGCTGCTTCTGTGT 1608
Qy      1738 GT--GGTGAGCAGAGTGGAGCTGGGGGGCGTGGAGAGCCGGGGCCCTGGCCACCTCCCT 1796
Db      1609 GTGGGTGATGGATGGAGCGGGGGCCCTGGAGAGCCCGTGGCCCTCCCACTCCCT 1668
Qy      1797 GACCC---GTCTAATATATAATATAGAGATGTGTCTATGG 1834
Db      1669 GTCCCTGTATCTAATATATAATATAGAGATGTGTATATGG 1709

```

RESULT 6

US-08-461-985-1

; Sequence 1, Application US/08461985

; Patent No. 5872006

; GENERAL INFORMATION:

; APPLICANT: Boulton, Teri G.

; APPLICANT: Cobb, Melanie H.

; APPLICANT: Yancopoulos, George D.

; APPLICANT: Nye, Steven

; APPLICANT: Panayotatos, Nikos

; TITLE OF INVENTION: A Family of Map2 Protein Kinases

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/461,985

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/176,620

; FILING DATE: 03-JAN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 6526-123

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-8864/9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1747 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1095

US-08-461-985-1

Query Match 55.8%; Score 1025; DB 2; Length 1747;
 Best Local Similarity 78.1%; Pred. No. 2.1e-202;
 Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

```

Qy      58 GAACCGAGGGGCTCGCCCGGGGTCCCGGGAGGTGGAGATGGTGAAGGGGCGAGCCGT 117
Db      5 GAATGCTGGGGTGTCTCCCGGTGTCTCCCGGGAGGTGGAGGTGGTGAAGGGGCGAGCAT 64
Qy      118 TCAGCTGGGGCCCGCGCTACACGAGTTGACGATCATCGCGAGGGCGGTACGGCATGG 177

```

```

Db      65 TCAGCTGGGGCCCGCTACACGCGCTGCAGTACATCGCGCGGGGCGGTACGGCATGG 124
Qy      178 TCAGCTGGGGCCCTATGACCAAGTGGCGCAAGATCTCGCGTGGGCCATCAAGAAGATCAGCCCCCT 237
Db      125 TCAGCTCAGCATATGACCAAGTGGCGCAAGATGGTGGTATCAAGAAGATCAGCCCCCT 184
Qy      238 TCAGCATCATCAGACCTACTGCGCAGCGCACCTCCGGGAGATCCAGATCTCTGCTGCGCTTC 297
Db      185 TCAGCATCAAAACCTACTGCTCAGCGCACCTGAGAGAAATCCAGATCTTGTCTCGGATTC 244
Qy      298 GCCATGAGAAATGTCAATCGGCATCCGAGACATCTTGGGGCGTCCACCTCGGAAGCCATGA 357
Db      245 GCCATGAGAAATGTCAATCGGCATCCGAGACATCTTCAAGACACCCACCTCGGAAGCCATGA 304
Qy      358 GAGATGTCTACATTTGTCGAGGACCTGATGAGAGATGACCTGTACAAAGTTGCTGAAAAGCC 417
Db      305 GAGATGTTTACATTTGTTCAAGGACCTCATGGAGACGACCTGTACAAAGTGTCTAAAGAGCC 364
Qy      418 AGCAGCTGAGCAATGACCATATCTGCTACTCTCTACAGATCTCTCGCGGGCTCAAGT 477
Db      365 AGCAGCTGAGCAATGACCATATCTGCTACTCTCTACAGATCTCTCGCGGGCTCAAGT 424
Qy      478 ACATCCACTCCGCAACGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACA 537
Db      425 ACATCACTCGGCCAATGTGCTGACCGGACCTGAAGCCCTCAATCTGCTTATCAACA 484
Qy      538 CCACCTGCGACCTTAAGATTTGATTTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 597
Db      485 CCACCTGCGACCTTAAGATTTGATTTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 544
Qy      598 ACCACCGGGCTTCTGAGCGGATGTGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
Db      545 ACCACCGGGCTTCTGAGCGGATGTGCTGACCGGATGTGCTGCTGCTGCTGCTGCTGCTGCT 604
Qy      658 TGCTGAACTCCAAAGGCTATACCAAGTCCATCGACATCTGGTCTGCTGCTGCTGCTGCTGCT 717
Db      605 TGCTTAACTCCAAAGGCTATACCAAGTCCATCGACATCTGGTCTGCTGCTGCTGCTGCTGCT 664
Qy      718 CTGAGATGTCTCTAAACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
Db      665 CTGAGATGTCTCTAAACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724
Qy      778 ACATTTCTGGGCTTCTGCGGCTCCCATCCAGAGGACCTGATTCGATTCATCAACATGA 837
Db      725 ACATTTCTAGGTATCTGCGGTTCCCATCCAGAGGACCTTAAATTTGATTCATTAACATGA 784
Qy      838 AGCCCGGAAACTACCTACAGTCTCTGCTCCCTCCAAAGACCAAGGTGGCTTGGGCCAAGCTTT 897
Db      785 AGCCCGGAAACTACCTACAGTCTCTGCTCCCTCTTAAACCAAGGTGGCTTGGGCCAAGCTTT 844
Qy      898 TCCCAAGTCAAGTCCAAAGCCCTTGACCTGCTGCAACCGGATGTTAAACCTTTAACCCCA 957
Db      845 TTTCCCAATCTGACTCCAAAGCTTTGACCTGCTGCAACCGGATGTTAAACCTTTAACCCCA 904
Qy      958 ATAAACGATCAAGTGGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
Db      905 ACAAGCGCATCAAGTAGAGGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
Qy      1018 CGACGATGAGCCAGTGGCGGCGGAGCCCTTACCTTTCGCGATGGAGCTGATGACCTAC 1077
Db      965 CGACAGATGAACCAAGTGGCTGAGAGGACCTTACCTTTGACATGGAGCTGATGATCTCC 1024
Qy      1078 CTAAAGGAGGGCTGAAGGAGCTCATCTTCCAGGAGACAGACGCTTCCAGCCCGGAGTGC 1137
Db      1025 CCAAGGAGCGGCTGAAGGAGCTGATCTTCCAAAGAGACAGCCGCTTCCAGCCAGGGGAC 1084
Qy      1138 TGAGAGGCGGCTTGAAGGAGACAGACATCTCTGCAACCTGGGGCTTGAAGAGTGGCA 1197
Db      1085 CAGAGGCGGCTTAAAGAAACAGACACCCCTGCTCTTTTGGACCTGG----- 1131
Qy      1198 AAGAGGCAAGAGGTCACTGAGGGCTCTGTGTCACCCAGGACCTGCTCTCTGCTGCTGCTC 1257
Db      1132 -----TCTGCTCTACCTGCTCTCT 1149

```

1258 TCCGCCAGACTGTTAGAAATGGACACTGTGCCAGCCGGACCTTGGACGCCAGGCC 1317
 1150 TCTCTGCAGATGTTAGAAATG--AACTTTGCTCAACCCGGACCCGGCAGCCAGG-C 1206
 1318 GGGGTGGAGCATGGGCTGGCCACCTCTCTCTTTGCTGAGGCTCCAGCTTTCAGGCAGG 1377
 1207 TGGACCAAGGTGGGCTGGACCCCTCTCACTCTGCTGGGTCTCTCTGTTCAAG--AG 1264
 1378 CCAAGGCTTCTCTCTCCCAACCGGCTCTCCCAAGGGGCTCGGGAGCTCAAGTGGCCCC 1437
 1265 GCTTCTCCCACTCCAGTCCCTGCCCCCATCTCCCTTGACCTGAGTGATGAGTGGTCCC 1324
 1438 AGTTCAATCTCCGCTGCTGCTGCGGCCCTTACCTTCCCGAGGTCCAGCTCTCTGGC 1497
 1325 AGAGCTGATCTCTGCTGCTG-TGCTTTATCTATCTCTGCTAGCCCGACCTCTGGTAGAC 1383
 1498 AGTTCTGGAATGAAGGGTCTGGCTGCCCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTG 1557
 1384 GGTCTGGAATGAAGGGCTATGACCGCCCTA-----GGACCTGTGCTACAGAGGGGTG 1437
 1558 GGGGGCGCTGAGTAGGACTCAGGGCCATGCTGCCCCCTCATCTCATTTCAAACCCCAAC 1617
 1438 GAGGGCACTGAGT-----AGGCTAAGCTCTGCCCCCTACTCATCTCTGTTGGAACCCCAAC 1489
 1618 CTTAGTTTCCCTGAAGGACATCTCTTACTCTCAAGGGCTAGCATCCCTGAGGAGCCAGG 1677
 1490 CCAATTTTCCCTGACAGAAATCTCTAAATCTCAAGGGCTAGTTTCCCTGAGGAGCCAGC 1549
 1678 CCGGGCCGAATCCCTCCCTGTCAAAAGCTGTCACTTCGCGGTGCGCTTCGTCTGTGT 1737
 1550 CTAGGCTTAACCTCTCCCTCTC-AAGCTGCCACATGTAACGCCCTTGTCTGTGT 1608
 1738 GT-GGTGAGCAGAGTGGAGCTGGGGCGTGGAGAGCCCGCGCCCTGCGCACTCCCT 1796
 1609 GTGGGTGATGGATGTGGAGCGGGCGCGTGGAGAGCCCGTGGCCCTCCCACTCCCT 1668
 1797 GACCC---GTCATATATATATATATAGAGATGTCTATGG 1834
 1669 GTGCCCTGTATCTAATATATATATAGAGATGTGTATATGG 1709

RESULT 7

US-08-458-887-1
 ; Sequence 1, Application US/08458887
 ; Patent No. 5914261
 ; GENERAL INFORMATION:
 ; APPLICANT: Boulton, Teri G. et al.
 ; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING: IBM PC compatible
 ; SOFTWARE: PC-DOS/MS-DOS
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/458,887
 ; FILING DATE: 02-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIORITY INFORMATION:
 ; APPLICATION NUMBER: US 07/701,544
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leblie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 6526-049

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 730-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1747 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1095
 ; US-08-458-887-1

Query Match 55.8%; Score 1025; DB 2; Length 1747;
 Best Local Similarity 78.1%; Pred. No. 2.1e-202;
 Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

58 GAAACCGAGGGGTGCGCCCGGGGTCCCGGGGAGGTGGAGATGTTGAAGGGGCGAGCCGT 117
 Db 5 GAACTGCTGGGGTCTGCTCCGGTGGTCCCGGGGAGTGGAGTGTGAAGGGGCGAGCCAT 64
 118 TCGACGTGGGCGCGCGCTACACGCAAGTTGCGAGTACATCGCGGAGGGCGCGTACGGCATGG 177
 Db 65 TCGAGCTGGGCCACGCTACACGCAAGTTGCGAGTACATCGCGGAGGGCGCGTACGGCATGG 124
 178 TCAGCTCGGCTTATGACCAAGTGGCGCAAGTTCGCGTGGCCATCAAGAAGATCAAGCCCT 237
 Db 125 TCAGCTCAGCATATGACCAAGTGGCGCAAGTTCGCGTGGCCATCAAGAAGATCAAGCCCT 184
 238 TCGAATCATGACCTTCTGCGGAGCGCGCTCCGGGAGTCCAGATCTCTGCTGCGCTCC 297
 Db 185 TCGAGCATCAAACTTCTGTCAGCGCACGCTGAGAGAAATCCAGATCTTCTGCGGATTC 244
 298 GCATGAGAAATGTCATCGGCATCCGAGATCTTCTCGGGCGTCCACCTCGGAAGCATGA 357
 Db 245 GCATGAGAAATGTCATGAGCATCCGAGATCTTCTCGGGAGCGCGCTCGGAAGCATGA 304
 358 GAGATGTCTACATTTGCGAGGACCTGATGGAGCTGACCTGTGTAAGTTGCTGAAAAGCC 417
 Db 305 GAGATGTTTACATTTCTCAGGACCTCATGGAGCGGACCTGTGTAAGTTGCTGAAAAGCC 364
 418 AGCAGCTGAGCAATGACCATATCTGCTACTTCTTACAGATCTTCTGCGGGCGCTCAAGT 477
 Db 365 AGCAGCTGAGCAATGACCATATCTGCTACTTCTTACAGATCTTCTGCGGGCGCTCAAGT 424
 478 ACATCCACTCCGCGCAACGTCCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACA 537
 Db 425 ACATACACTCGGCCAATGTGCTGACCGGACCTTGAAGCCCTCCAATCTGCTTATCAACA 484
 538 CCACCTGGACCTTAAAGATTTGATTTTCGGCTGGCCCGGATTCGCCGATCTCTGAGCATG 597
 Db 485 CCACCTGGACCTTAAAGATTTGATTTTCGGCTGGCCCGGATTCGCCGATCTCTGAGCATG 544
 598 ACCACACCGGCTTCTGCGGAGTATGTGGCTGACCGGTGTTACCGGGCCCGAGAGATCA 657
 Db 545 ACCACACCGGCTTCTGCGGAGTATGTGGCTGACCGGTGTTACCGGGCCCGAGAGATCA 604
 658 TGCTGAACTCCAAAGGGCTATACCAAGTCCATCGACATCTGTGTTGGGCTGCATTTCTGG 717
 Db 605 TGCTTAACTCCAAAGGGCTATACCAAGTCCATCGACATCTGTGTTGGGCTGCATTTCTGG 664
 718 CTGAGATGCTCTTAAACCGGCCCATCTTCCCTGGCAAGCATCTACCTGGATGAGCTCAACC 777
 Db 665 CTGAGATGCTCTTAAACCGGCCCATCTTCCCTGGCAAGCATCTACCTGGATGAGCTCAACC 724
 778 ACATTTCTGGCATCTCTGGGCTCCCATCCAGAGGAGCTGAAATGTTATCATCAACATCA 837
 Db 725 ACATTTCTGGTATCTCTGGGTTCCCATCCAGAGGAGCTGAAATGTTATCATTAACATGA 784
 838 AGGCCCGAAACTTACCTTACAGTCTCTGCGCTCCAAAGCAAGGTGGCTTGGGCCAAGCTTT 897

785	Db	 AGCCCCGAACTACCTACAGTCTCTGCGCTCTAAAAACCAGAGTGGGTGGCGCCAAAGCTTT	844
898	Qy	 TCCCCAAGTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAAACCTTTAAACCCCA	957
845	Db	 TTCCCAAACTGACTCCAAAGCTCTTGACCTGCTGGACCGGATGTTAAACCTTTAAACCCAA	904
958	Qy	 ATAAACGATCA CAGTGGAGGAAGCGTGTGGCTCACCCCTACCTTGGAGCAGTACTATGACC	1017
905	Db	 ACAAGGGCATCA CAGTAGAGGAAGCACTGGCTCACCCCTTACCTTGGAAACAGTACTATGATC	964
1018	Qy	 CGACGATGAGCCAGTGGCGCGAGAGCCCTTACCTTGGCCATGGAGCTGGATGAACCTAC	1077
965	Db	 CGACAGATGAAACAGTGGCTTGAGGAGCCATTACCTTTGACATGGAGCTGGATGATCTCC	1024
1078	Qy	 CTAAGGAGCGGCTGAAGGAGCTCATCTTCAGAGAGACAGCACGCTTCCAGCCCGGAGTGC	1137
1025	Db	 CCAAGGAGCGGCTGAAGGAGCTCATCTTCCAAAGAGACAGCCGCTTCCAGCCAGGGGCGAC	1084
1138	Qy	 TGGAGGCCCTTAGCCAGACAGACATCTCTGCACCTCGGGGCTGTGAAACAGAACTGGCA	1197
1085	Db	 CAGAGGCCCTTAACAAGAAACAGACACCCCTGTCTTTTGGACTGG-----	1131
1198	Qy	 AAGAGGCNAGAGTCACTGAGGGCCTTGTGCA CCGAGGACCTGCTCTGCTGCCCCCTC	1257
1132	Db	 -----TCTGCTCTACCTGCTCCT	1149
1258	Qy	 TCCGGCCAGACTGTTAGAAAATGGACACTGTGCGCCAGCCGAGACCTTGGCAGCCCAAGGC	1317
1150	Db	 TCTCTGCAGATTGTTAGAAAATG--AACTTTGTCTAACCCGAGACCCCGGAGCCCAAGG-C	1206
1318	Qy	 GGGGTGAGCATGGGCTCGGCCACTCTCTCTCTTGTCTGAGGGCTCCAGCTTTCAGGCAGG	1377
1207	Db	 TGAAACNAGGGTGGGCTTGSCACCCCTCTCACTCTGCTGGGGTCTCTCTGTTCAAG--AG	1264
1378	Qy	 CCAAGGCCCTTCTCTCCCCACCCGCGCTTCCCA CCGGGGCTTCGGGAGCTCAGGTGGCCCC	1437
1265	Db	 GCTTCTCCCACTCCAGTCCGCTGCGCCCATCTCCCGCTTGACCTGAGTGATGAGGTGGTCCC	1324
1438	Qy	 AGTTCAATCTCCCGCTGCTGCTGCGCGCTTACCTTCCCGAGCGTCCAGTCTCTGGC	1497
1325	Db	 AGAGCTGATCTCTGCTGCTG--TGTTCTTTATCTATCCCTGCTAGCCCCAGCTCTGTTAGAC	1383
1498	Qy	 AGTTCTTGGAAATGAAGGGTTCTGGCTGCCCCAAACCTGCTCAAGGGCAGAGGTGAGGGTG	1557
1384	Db	 GGTTCTTGGAAATGAAGGGCTATGACGCCCTA-----GGACCTGTGCTACAGAGGGGTG	1437
1558	Qy	 GGGGGCGCTGAGTAGGAGACTCAGGGCCATGCTGCCCGCTCATCTCATTTCAAAACCCAC	1617
1438	Db	 GAGGGCACTGAGT-----AGGCTAAGCTCTGCCCTACTCATCTCTGTTGGAAACCCAC	1489
1618	Qy	 CCTAGTTTCCCTGAAGGAACATTCCTTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGG	1677
1490	Db	 CCCATTTTCCCTTGACAGAACATTCCTAAATCTCAAGGGCTAGTTTCCCTGAGGAGCCAGC	1549
1678	Qy	 CCGGGCGGAATCCCTCCTGTCNAAAGCTGTCACTTCGGTGGCTCGCTGCTCTCTGTGT	1737
1550	Db	 CTAGGCCATAACCCCTCTCCCTCTC--AAGCTTGCCACATGTAAACGCCCTTGTGCTTCTGTGT	1608
1738	Qy	 GT--GGTCAGCAGAAAGTGGAGCTGGGGGGCGTGGAGAGCCCGGGCGCCCTGCCACCTCCCT	1796
1609	Db	 GTGGGTGATTGGATGTGGAGCGGGGGCCGTTGGAGAGCCCGTGGCCCCCTCCCACTCCCT	1668
1797	Qy	 GACCC---GTCTAATATATAAATAGAGATGTGCTATGG	1834
1669	Db	 GTGCTGTATCTAATATATAAATAGAGATGTGTATATGG	1709

RESULT 8

US-08-932-787B-1

US-08-932-787B-1
: Sequence 1, Application US/08932787B

; Patent No. 6277963

FACEID NO. 0277205
; GENERAL INFORMATION:

QY 718 CTGAGATGCTCTTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGTCAACCC 777
Db 665 CTGAGATGCTCTTAAACCGGCCCATCTTCCCGGCAAGCACTACCTGGACGACTCAACC 724
QY 778 ACATTCCTGGGCATCTCTGGGCTCCCATCCAGAGAGGACCTGAATTGTATCATCAACATGA 837
Db 725 ACATTCAGGTATCTACTGGGTTCGCCATCCCAAGAGGACCTAAATTGTATCATTAACATGA 784
QY 838 AGSCCGGAACCTACCTACAGTCTCTGCCCTTCCAGACCAAGTGGCTTGGGCAAGCTTT 897
Db 785 AGSCCGGAACCTACCTACAGTCTCTGCCCTTAAACCAAGTGGCTTGGGCAAGCTTT 844
QY 898 TCCCCAAGTCAGACTCCAAAGCCCTTGACCTGTCTGAGCCGAGTGTAACTTTAAACCCCA 957
Db 845 TTCCCCAATCTGACTCCAAAGCTCTTGACCTGTCTGAGCCGAGTGTAACTTTAAACCCCA 904
QY 958 ATAAACGGATCAGTGGAGGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACC 1017
Db 905 ACAGCGCATCACAGTAGAGGAAGCACTGGCTCACCCCTTACCTGGAAACAGTACTATGATC 964
QY 1018 CGACGATGAGCCAGTGGCGGAGGAGCCCTTACCTTCCGCATGAGCTGGATGACCTTAC 1077
Db 965 CGACATCAACCCAGTGGCTGAGGAGCCATTCACCTTTGACATGAGCTGGATGATCTCC 1024
QY 1078 CTAAGGAGCGGCTGAAGGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGC 1137
Db 1025 CCAAGGAGCGGCTGAAGGAGCTGATCTTCCAGAGACAGCCGCTTCCAGCCAGGAGCAC 1084
QY 1138 TGAGGCGCCCTAGCCGACAGACAGACATCTCTGCACCCCTGGGCGCTGGAACTGGCA 1197
Db 1085 CAGAGCGCCCTAACAAAGAACAGACACCCCTGTCTTTTGGACCTGG----- 1131
QY 1198 AAGAGCAAGAGGTCACTGAGGCGCTCTGTACCCAGGACCTGCTCTCTGCGCTGCCCTTC 1257
Db 1132 -----TCTGCTCTACTGCTCTCT 1149
QY 1258 TCCGCGCAGACTGTTAGAAAATGGACACTGTGCGCAGCCCGGACCTTGGCAGCCAGAGCC 1317
Db 1150 TCTCTGCAATGTTTGAANAATG--AACTTTGTCTAACCCGAGACCCCGGACGCCAAG-C 1206
QY 1318 GGGGTGGAGCATGGGCTGGCCACCTCTCTCTTTTGTCTGAGGCGCTCCAGCTTTCAGCAG 1377
Db 1207 TGGACCAAGGTGGGCTGGCACTCTCTCACTGTGCTGGGTCTCTCTGTTCAAG--AG 1264
QY 1378 CCAAGGCCCTTCTCTCCCAACCGGCCCTCCCAAGGGGCTCTGGAGCTCAGGTGGCCCC 1437
Db 1265 GCTTCTCCCACTCCAGTCCCGCTGCCCATCTCCCTTTGACCTGAGTGAATGAGTGGTCCC 1324
QY 1438 AGTTCAATCTCCGCTGCTGCTGGCGCCCTTACCTTCCCGAGCTCCAGCTCTCTGGC 1497
Db 1325 AGAGCTGATCTCTGCTGCTG-TGTCTTTATCTATCTCTCTGACCCCACTCTGTTAGAC 1383
QY 1498 AGTTCTGGAATGGAAGGGTCTGGCTGCGCCCAACCTGTCTGAAGGGCAGAGGTGGAGGGTG 1557
Db 1384 GGTTCGGAATGGAAGGGTATGACCGCCCTA-----GGACTGTGCTACAGAGGGGTG 1437
QY 1558 GGGGGCGCTGAGTAGGAGCTCAGGGGCATGCTCTGCCCCCTCATCTCATTTCAAAACCCCA 1617
Db 1438 GAGGGCACTGAGT-----AGGCTAAGCTCTGCCCCCTACTCATCTGTGTTGAAACCCCA 1489
QY 1618 CCTAGTTTCCCTGGAAGAACNTTCCTAGTCTCAAGGGCTAGCANTCCCTGAGGAGCCAGG 1677
Db 1490 CCAATTTTCCCTGACAGAACATTCCTAAATCTCAAGGGCTAGTTTCCCTGAGAGGACGAC 1549
QY 1678 CCGGGCGGAATCCCTCTCCCTGTCAAAAGCTGTCACTTCGCGTGGCCCTCGCTCTGTGT 1737
Db 1550 CTAGGCTTAACCTCTCTCCCTCTC-AAGCTGCCACATGTAAGCCCTTGTCTGTCTGTGT 1608
QY 1738 GT-GGTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCCGGGCGCCCTTGCACCTCCCT 1796
Db 1609 GTGGGTGATTGGATGTGGAGGCGGGCGCGTGGAGAGCCCGTGGAGAGCCCGTCCCTCCCT 1668

QY 1797 GACCC--GTCTAATATATAAATATAGAGATGTGTCTATGG 1834
Db 1669 GTGCGTGTATCTAATATATAAATATAGAGATGTGTATATGG 1709
RESULT 9
US-08-932-012C-1
; Sequence 1, Application US/08932012C
; Patent No. 6297035
; GENERAL INFORMATION:
; APPLICANT: Boulton et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; FILE REFERENCE: REG 430-Y-1
; CURRENT APPLICATION NUMBER: US/08/932,012C
; CURRENT FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 08/462,874
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 08/178,488
; PRIOR FILING DATE: 1994-01-07
; PRIOR APPLICATION NUMBER: 07/701,544
; PRIOR FILING DATE: 1991-05-16
; PRIOR APPLICATION NUMBER: 07/532,004
; PRIOR FILING DATE: 1990-06-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1747
; TYPE: DNA
; ORGANISM: RAT
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1095)
; OTHER INFORMATION: ERK1 cDNA
US-08-932-012C-1
Query Match 55.8%; Score 1025; DB 3; Length 1747;
Best Local Similarity 78.1%; Pred. No. 2.1e-202;
Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;
QY 58 GNAACGAGGGGTGCGCGCGGGTCCGCGGGAGGTGGAGATGTGAAGGGCAGCCGT 117
Db 5 GAACTGCTGGGTGCTGCTCCGGTGGTCCCGGGAGGTGGAGTGTGAAGGGCAGCCAT 64
QY 118 TCGACGTGGCGCGCGCTACACGACGATTTGCAGTATCATCGCGAGGGCGGTACGGCATGG 177
Db 65 TCGAGTGGGCCACGCTACACGACGCTGACGATCATCGCGAGGGCGGTACGGCATGG 124
QY 178 TCAGTTCGGCTTATGACCACTGTCGCAAGACTGCGGTGGCCATCAAGAGATCAGCCCT 237
Db 125 TCAGCTCAGCATATGACCACTGTCGCAAGACTGCGGTGGCCATCAAGAGATCAGCCCT 184
QY 238 TCGAACATCAGACCTACTGCGCAGCGCATCGCGGAGATCCAGATCTTGGCTGGCTTCC 297
Db 185 TCGAGCATCAAACTACTGTCAGCGCACGCTGAGAGAAATCCAGATCTTGGCTGGATTC 244
QY 298 GGCATGAGATGCTCATCGGCATCCGAGACATTTCTCGGCGCTCCACCTTGAAGCCATGA 357
Db 245 GGCATGAGATGCTCATAGGCATCCGAGACATCTCAGAGCACCCCTTGAAGCCATGA 304
QY 358 GAGATGCTTACATTTGCGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAAGCC 417
Db 305 GAGATGTTTACATTTGTCAGGACCTCATGGAGACCTCATGGAGACGCACTGTACAAGCTGCTTAAAGAGCC 364
QY 418 AGCAGCTGAGCAATGACCAATATCTCTACTTCTCTACAGATCTCTGGGGCCCTCAAGT 477
Db 365 AGCAGCTGAGCAATGACCAATCTCTCTACTTCTCTACAGATCTCTGGGGCCCTCAAGT 424
QY 478 ACATCCACTCCGCCAAAGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGTCTCATCAACA 537
Db 425 ACATCACTCCGCCAAATGTGCTGCAACCGGACCTGAGCCCTCCAACTGTCTCATCAACA 484
QY 538 CCACCTGCGACCTTAAGATTTTGTGATTTTGGCGCTGGCCCGGATTCGCGATCTCTGAGCATG 597

Db	485	CCACCTGCGACCTTAAGATCTGTGATTTTGGCCCTTGGCCCGATGTCGACCCCTGAGCAG	544
Qy	598	ACCACACCGGCTTCTGACCGAGTATGTGGCTACGGCTGCTGACCGGCGCCCGACGATCA	657
Db	545	ACCACACTGGCTTCTGACCGAGTATGTGGCCACACGCTGATACCGAGCCCGACGATCA	604
Qy	658	TGCTGAATCTCAAGGGCTATACAGAGTTCATCGACATCTGGTCTGTGGGCTGCAATCTGG	717
Db	605	TGCTTAATCTCAAGGGCTACACCAATTCATTCGACATCTGGTCTGTGGGCTGCAATCTGG	664
Qy	718	CTGAGATGCTCTTAACCGGCCCATCTTCCCTGCGACGACTACCTGGATCAGCTCAACC	777
Db	665	CTGAGATGCTCTCAACCGGCCCATCTTCCCGCGACGACTACCTGGACAGCTCAACC	724
Qy	778	ACATTTGGGCGATCTGGGCTCCCATCCAGGAGGACCTGAAATTTGTATCATCAACATGA	837
Db	725	ACATTTAGGTATCTGGGTTCCCATCCCAAGGAGCTAAATTTGTATCATTAACATGA	784
Qy	838	AGGCCGGAATCTACCTACAGTCTTGCCTTCCAGAACCAAGGTGGCTTGGGCCCAAGCTTT	897
Db	785	AGGCCGGAATCTACCTACAGTCTTGCCTTCCAAACCAAGGTGGCTTGGGCCCAAGCTTT	844
Qy	898	TCCCCAGTCAAGTCCAAAGCCCTTGAACCTGTGACCGGATGTTAACTTTAAACCCCA	957
Db	845	TTCCCAATCTGACTCCAAAGCTCTTGACCTGTGAGCCGATGTTAACTTTAAACCCCA	904
Qy	958	ATAACCGATCACAGTGGAGAGCGCTGGCTCACCCCTACCTGGAGCAGTACTATGACC	1017
Db	905	ACHAGCCATCACATGAGAGGACATGGCTCACCTTACCTGGACAGTACTATGATC	964
Qy	1018	CGACGGATGACCGACGTCGCGAGAGCCCTTCACTTCCGATCGGAGCTGGATGACCTAC	1077
Db	965	CGACAGATGAACAGTGGCTGAGGAGCAATTCACCTTTGACATGGAGCTGGATGATCTCC	1024
Qy	1078	CTAAGGCGGCTCAAGAGCTCATCTTCCAGAGACAGACGCTTCCAGCCCGGAGTGC	1137
Db	1025	CCAAGGAGCGCTGAAGAGCTGATCTTCAAGAGACAGCCCGCTTCCAGCCCGGAGC	1084
Qy	1138	TGAGGCGCCCTAGCCACAGACAGACATCTCTGACCCCTGGGCGCTCGAAACAGAACTGGCA	1197
Db	1085	CAGGCGCCCTTACACAGACAGACACCCCTGTCTTTTGGACCTGG-----	1131
Qy	1198	AAGGCGAAGAGTCACTGAGGGCTCTGTACCCAGGACCTGCTCTCTGCTGCGCCCTC	1257
Db	1132	-----TCTGCTCTACTGCTCTCT	1149
Qy	1258	TCCCGCAGACTGTTAGAAAATGACACTGTGCCAGCCCGGACCTTGGCAGCCAGGCC	1317
Db	1150	TCTCTGAGATTTGTTAGAAAATG--AACTTTGTCTCAACCCCGGACCCCGGAGCCAGG-C	1206
Qy	1318	GGGCTGGAGCATGGCCCTGGCCACCTCTCTCTTTGTGAGGCTTCCAGCTTCAAGCAGG	1377
Db	1207	TGGAACCAAGGGTGGGCTTGGACCCCTCTCACTGTGTGGGTCTCTCTGTTCAAG--AG	1264
Qy	1378	CCAAGGCTTCTCTCCCTCCACCCGCTTCCCAAGGGGCTTCCGAGCTCAGGTGGCCCC	1437
Db	1265	GCTTCTCCACTCCAGTCCCTGCCCCATCTCCCTTTGACCTGATGATGAGGTGGTCCC	1324
Qy	1438	AGTTCAATCTCCCTGCTGTGCTGTGCGCCCTTACCTTCCCGAGCTGCCAGTCTCTGGC	1497
Db	1325	AGAGCTGATCTCTGCTGTG--TGTCTTTATCTATCTCTCTGCTAGCCCGAGCTCTGGTAGAC	1383
Qy	1498	AGTTCTGAATGGAAGGTTCTGCTGCGCCCAACCTGTCTGAAGGCGCAGAGTGGAGGGTG	1557
Db	1384	GTTTCTGAATGGAAGGCTATGACCGCCCTA-----GGACCTGTGCTACAGAGGGGTG	1437
Qy	1558	GGGGCGCTGAGTAGGGACTCAGGGCCATGCTCTCCCGCTCTCATCTCAATCAAAACCCAC	1617
Db	1438	GAGGGCACTGAGT-----AGGCTAAGCTCTGCCCTTACTCATCTCTTGGAAACCCAC	1489
Qy	1618	CCTAGTTTCCCTGAAGGAACATTCCTTAGTCTCAAGGGCTAGCATCTCTGAGGAGCCAGG	1677
Db	1490	CCCATTTTCCCTGACAGAAATTCCTAATCTCAAGGGCTAGTTTTCCTTGAGGAGCCAGC	1549

Qy	1678	CGGGCGCGAATCCCTTCCCTGTCAAAGCTGTCACTTCGGCTGCCCTCGCTTCTGTGT	1737
Db	1550	CTAGGCGCTAACCTCTCCCTCTC-AACTGCCCATGTAAAGCCCTTGTCTTCTGTGT	1608
Qy	1738	GT--GGTGACAGAAAGTGGAGCTGGGGGGCTGGAGAGCCGGGGCCCTTGCACCTCCCT	1796
Db	1609	GTGGGTGATTTGATGTGGAGGGGGCCCTGTGGAGAGCCCGTGCCTCCCACTCCCT	1668
Qy	1797	GACCC---GTCTAATATATAAATATATAGAGATGTCTATGG	1834
Db	1669	GTGCCTGTATCTAATATATAAATATATAGAGATGTGTATATGG	1709

RESULT 10
US-08-888-818C-1
; Sequence 1, Application US/08888818C
; Patent No. 6303358
; GENERAL INFORMATION:
; APPLICANT: Boulton et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; FILE REFERENCE: REG 430-V-1
; CURRENT APPLICATION NUMBER: US/08/888,818C
; CURRENT FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: 08/478,985
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/178,488
; PRIOR FILING DATE: 1994-01-07
; PRIOR APPLICATION NUMBER: 07/701,544
; PRIOR FILING DATE: 1991-05-16
; PRIOR APPLICATION NUMBER: 07/532,004
; PRIOR FILING DATE: 1990-06-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1747
; TYPE: DNA
; ORGANISM: RAT
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1095)
; OTHER INFORMATION: ERK1 cDNA
US-08-888-818C-1

Query Match 55.8%; Score 1025; DB 3; Length 1747;
Best Local Similarity 78.1%; Pred. No. 2.1e-202;
Matches 1391; Conservative 0; Mismatches 310; Indels 80; Gaps 10;

Qy	58	GAACCGAGGGGTTCGGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCGT	117
Db	5	GAACCTGTGGGGTCTGTCCTCCGGTGGTCCCGGGGAGGTGGAGGTGAAGGGGCGAGCAT	64
Qy	118	TCGACGTGGGCGCCGCTACACGAGTTGCAGTACATCGCGGAGGGCGGTACGGCATGG	177
Db	65	TCGACGTGGGCGCCGCTACACGAGTTGCAGTACATCGCGGAGGGGCGGTACGGCATGG	124
Qy	178	TCAGCTCGGCTATGACCACTGCGCAAGACTCGCTGGCCATCAAGAGATCAGCCCT	237
Db	125	TCAGCTCAGCATATGACCACTGCGCAAGACTCAGAGTGGCTATCAAGAGATCAGCCCT	184
Qy	238	TCGAATCAGACTTACTGCGGAGGACGCTCCGGGAGATCCAGATCTCTGCGCTTCC	297
Db	185	TCGAGCATCAACCTTACTGTGCGCACGCTGAGAGAAATCCAGATCTTCTCGATTCC	244
Qy	298	GCCATGAGATGTTCATCGGATCCGAGACATCTTCGGGGCGTCCACCTCGAAGCCATGA	357
Db	245	GCCATGAGATGTTCATAGGATCCGAGACATCTTCAGAGCACCCCTCGAAGCCATGA	304
Qy	358	GAGATGTCTACATTTGTGAGGACCTGATGGAGACTGACCTGTACAAGTTGCTGAAAGCC	417
Db	305	GAGATGTTTACATTTTTCAGGACCTCATGAGAGCGGACCTGTACAAGCTGTCTAAGAGCC	364
Qy	418	AGCAGCTGAGCAATGACCATATCTGCTACTTCTTACCAGATCTCTCGGGGCGCTCAAGT	477

Db 365 |||||AGCAGCTGAGCAATGACCACATCTGCTACTTCTCTTACAGATCCTCCGGGCTCAAGT 424
Qy 478 ACATCACTCCGCGCAACGCTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACA 537
Db 425 ACATCACTCGGCAATGTGCTGCACCGGACCTGAAGCCCTCCAATCTGCTTATCAACA 484
Qy 538 CCACCTGCGACCTTAAGATTGTGATTTGGGCTGGCCCGGATTCGGATCCGTGAGCATG 597
Db 485 CCACCTGCGACCTTAAGATTGTGATTTGGGCTTCGCCGGATGTGTGACCTGTGACGACG 544
Qy 598 ACCACACCGGCTTCCTGACCGAGTATGTGGCTACGCGCTGTGTACCGGGCCCGAGAGATCA 657
Db 545 ACCACACTGGCTTCTGACCGAGTATGTGGCCACACGCTGTGTACCGAGCCCGAGAGATCA 604
Qy 658 TGTGTAATCTCAAGGGCTATACCAAGTCCATCGACATCTGTGTCTGTGGGCTGTGATCTGG 717
Db 605 TGTCTAACTCCAAGGGCTACACCAATCCATTGACATCTGTGTGTGGGCTGTGATCTGG 664
Qy 718 CTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGTCAAC 777
Db 665 CTGAGATGCTCTCAACCGGCCATCTTCCCGGCAAGCACTACCTGGACCAAGCTCAAC 724
Qy 778 ACATTTCTGGGCATCTCTGGGCTCCCATCCAGGAGCACTGTAATTGTATCATCAACATGA 837
Db 725 ACATTTCTAGGTACTGGGTTCCCATCCCAAGAGGACCTAAATTGTATCATTAACATGA 784
Qy 838 AGCCCGCAATCACTACAGTCTCTGCCCTCCAGACCAAGTGGCTTGGGCGCAAGCTTT 897
Db 785 AGCCCGCAATCACTACAGTCTCTGCCCTTAAACCAAGTGGCTTGGGCGCAAGCTTT 844
Qy 898 TCCCCAAGTCAGACTCCAAAGCCTTGACTGCTGACCGGATGTAACTTTAAACCCCA 957
Db 845 TTTCCCAAACTGACTCCAAAGCTCTTGACTGCTGACCGGATGTAACTTTAAACCCCA 904
Qy 958 ATAAACGGATCAGTGGGGAAGCGCTGGCTCACCCCTACTTGGAGCAGTACTATGACC 1017
Db 905 ACNAGGCAATCAGTAGAGGAGCACTGGCTCACCTTACCTGGNACAGTACTATGATC 964
Qy 1018 CGACGATGAGCGAGTGGCGGAGGAGCCCTTCACTTTCGCGCATGTGAGCTGGATGACTAC 1077
Db 965 CGACAGATGAACAGTGGCTGAGGAGCCATTACCTTTGACATGGAGCTGGATGATCTCC 1024
Qy 1078 CTAAAGAGCGGCTGAGGAGCTCATCTTCCAGGAGCAGCAGCTTCCAGCCGCGAGTGC 1137
Db 1025 CCAAGGAGCGGCTGAAGGAGCTGATCTTCAAGAGACAGCCCGCTTCCAGCCAGGGGAC 1084
Qy 1138 TGAGGCGCCCTAGCCCGACAGACATCTCTGACCCCTGGGCGCTTGGAAACAGAACTGGCA 1197
Db 1085 CAGAGGCCCTTAAAGAACAGACACCCCTGTCTTTTGGACCTGG-----1131
Qy 1198 AAGAGCAAGAGTCACTGAGGCGCTCTGTGACCCAGGACCTGCTCTGCTGCTGCCCTC 1257
Db 1132 -----TCTGCTCTACCTGCTCCT 1149
Qy 1258 TCCCGCCAGACTGTTAGAAAATGGACACTGTGCGCCAGCCCGGACCTTGTGCGCCAGGCG 1317
Db 1150 TCTCTGAGATTGTTAGAAAATG--AACTTTGTCTAAACCGGACCCCGGAGCCCAAGG-C 1206
Qy 1318 GGGGTGGAGATGGGCGCTGGCCACCTCTCTCTCTTTGTCTGAGGCGCTCCAGCTTCAGG 1377
Db 1207 TGGACCAAGGTGGGCGCTGGCAACCTCTCACTGTGTGGGCTCTCTCGTTCAAG--AG 1264
Qy 1378 CCNAGGCTTCTCTCCCAACCGGCGCTCCCAAGGGGCGCTGGGAGCTCAGGTGGCGCC 1437
Db 1265 GCTTCTCCCACTCCAGTCCCTCTGCCCATCTCCCTTTGACCTGAGTGTGATGGGTGGTCC 1324
Qy 1438 AGTTCAATCTCCGCTGCTGCTGTGCGCCTTTACCTTCCCAAGCGTCCCACTCTGTGC 1497
Db 1325 AGAGCTGATCTGTGCTGTG--TGTCTTTATCTATCCCTGTGTAGCCCAAGCTCTGTGTAG 1383
Qy 1498 AGTTCTGGAAATGGAAAGGTTCTGGCTGCCCAACCTGCTGAAGGGCAGAGGTGGAGGTG 1557
|||||

Db 1384 GGTTCTGGAAATGGAAAGGCTATGACGCCCTA-----GGACCTGTGTCTACAGAGGGGTG 1437
Qy 1558 GGGGGCGCTGAGTAGGGACTCAGGGCCATGCTGCTGCCCTCATCTCATTTCAAAACCCAC 1617
Db 1438 GAGGCGACTGAGT-----AGGCTAAGCTCTGCCCTACTCATCTGTTGAAACCCAC 1489
Qy 1618 CCTAGTTTCCCTGGAAGGAACATTCTTATAGTCTCAAGGGCTAGCATCCCTTGAGAGCCAG 1677
Db 1490 CCATTTTCCCTGACAGAACATTCTTAATCTCAAGGGCTAGTTTCCCTTGAGAGCCAG 1549
Qy 1678 CCGGCGCGAATCCCTCCCTGTCAAAGTGTCACTTTCGGGTGCCCTCGCTGCTTCTGTGT 1737
Db 1550 CTAGGCTTAACCTCTCCCTCTC-AAAGTGCACATGTAAACGCCCTTGTGCTTCTGTGT 1608
Qy 1738 GT-GGTGACAGAACTGGAGCTGGGGGGGTGGAGAGCCCGGCCCTTGCACCTCCCT 1796
Db 1609 GTGGGTGATTTGATGTGGAGGGGGCGGTGGAGAGCCCGTGGCCCTCCCCACCTCCCT 1668
Qy 1797 GACCC---GTCTTAATATATAAATATAGAGATGTGTCTATGG 1834
Db 1669 GTGCTGTATCTTAATATATAAATATAGAGATGTGTATATGG 1709

RESULT 11
US-09-417-197-56
; Sequence 56, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mERK1-F64L-S65T-GFP fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1872)
US-09-417-197-56

Query Match 52.2%; Score 958.6; DB 3; Length 1875;
Best Local Similarity 90.0%; Pred. No. 1.1e-188;
Matches 1027; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 14 GCGCGCGGCGGCGGCTCAGGGGGCGGGGGCGGGGAGCCCGTAGAACCGAGGGGTGCG 73
Db 6 GCGCGCGGCGGCGGCTCAGGGGGCGGGGGCGGGGAGCCCGAGGGAACTGCTGGGGTCTG 65

Qy 74 CCGCGGGGTCCCGGGGAGGTGAGATGCTGAAGGGGAGCGGCTTCGACGTGGGCGCGCG 133
Db 66 CCGGTGTGTCCTCGGGGAGGTGAGGTGTGAAGGGGAGCCATTTCGATGTGGGCCACG 125

Qy 134 CTACACGCAATTTGCAGTACATCGGCGAGGCGCGCTACGGCATGGTCAGCTCGGCTATGA 193
Db 126 CTACACGCACTGTCAGTACATCGGCGAGGCGCGCTACGGCATGGTCAGCTCAGCTTATGA 185

Qy 194 CCAGTGGCGACACTCCGCTGGCCATCAAGAAGATCAGCCCTTCGACATCAGACCTA 253
Db 186 CCAGTGGCGAAGACAGAGTGGCCATCAAGAAGATCAGCCCTTTGAGCATCAAACTA 245

Qy 254 CTGCCAGCGCAGCTCCCGGGAGATCCAGATCTGTGCTGCCCTTCGCGCATGAGAAATGAT 313
Db 246 CTGTCAAGCGACGCTGAGGGAGATCCAGATCTTGTGCGATTCGCGCATGAGAAATGAT 305

Qy 314 CCGCATCCGAGACATTTCTGGGGGCGTCCACCTCGAAGCCATGAGAGATGTCTACATTGT 373
|||||

306 AGGCATCCGAGACATCTCCAGAGCGCCACCCCTGGAAGCATGAGAGATGTTTACATTTGT 365
Db
374 GCAGGACCTGATGAGACTGACCTGTACAAAGTTGCTGAAAGCCAGCAGCTGAGCAATGA 433
Qy
366 TCAGGACCTCATGGAGACAGACTGTACAGCTGCTTAAAGCCAGCAGCTGAGCAATGA 425
Db
434 CCATATCTGCTACTTCTCTACAGATCTCTGCGGGGCTCAAGTATACATCCACTTCGCGCAA 493
Qy
426 CCACATCTGCTACTTCTCTACAGATCTCTGCGGGGCTCAAGTATATATACATCAGCCAA 485
Db
494 CTTGCTCCAGGAGATCTAAAGCCCTCAACCTGCTCATCAACACACCTGCGACCTTAA 553
Qy
486 TGTGCTGACCGGACCTGAAGCTTCCAACTGCTTATCAACACACCTGCGACCTTAA 545
Db
554 GATTGTGATTTGGGCTGCGCGGATTCGCGATCTCTGAGCATGACACACCGGCTTCCT 613
Qy
546 GATCTGTGATTTGGGCTGCGCGGATTCGCGATCTCTGAGCATGACACACCTGCGCTTCT 605
Db
614 GACGAGTATGTGGCTGAGCTGATACCGGGGCTCCAGAGATCATGCTGAACTCCAAAGG 673
Qy
606 GACGAGTATGTGGCTGAGCTGATACCGGGGCTCCAGAGATCATGCTGAACTCCAAAGG 665
Db
674 CTATACCAAGTCCATGACATCTGCTGTGGGCTGCAATCTGCTGAGATGCTCTCTAA 733
Qy
666 CTACACCAATCCATGACATCTGCTGTGGGCTGCAATCTGCTGAGATGCTCTCTCAA 725
Db
734 CCGGCCCATCTTCCCTGGCAAGCACTACCTGGATGCTCAACACCATCTTGGGCTACCT 793
Qy
726 CCGGCCCATCTTCCCTGGCAAGCACTACCTGGATGCTCAACACCATCTTGGGCTACCT 785
Db
794 GGGCTCCCATCTCCAGGAGCACTGAAATGTATCATCAACATGAAGCCCGGAACTACCT 853
Qy
786 GGGTTCCTCCATCCAGGAGCACTTAAATGTATCATCAATTAACATGAAGCCCGGAACTACCT 845
Db
854 ACAGTCTCTGCTCCAGGAGCACTGAAATGTATCATCAACATGAAGCCCGGAACTACCT 913
Qy
846 GCGTCTCTGCTCCAGGAGCACTTAAATGTATCATCAATTAACATGAAGCCCGGAACTACCT 905
Db
914 CAAGGCTCTGCTCCAGGAGCACTGAAATGTATCATCAACATGAAGCCCGGAACTACCT 973
Qy
906 CAAGGCTCTGCTCCAGGAGCACTTAAATGTATCATCAATTAACATGAAGCCCGGAACTACCT 965
Db
974 GGAGGAGCGCTGGCTCACCCCTACCTGAGCAGTACTATGACCCGAGGATGAGCCAGT 1033
Qy
966 AGAGGAGCGCTGGCTCACCCCTACCTGAGCAGTACTATGACCCGAGGATGAGCCAGT 1025
Db
1034 GGCGGAGGCGCTTACCTTGGCCATGAGCTGAGTACCTTACCTAAGGAGCGGCTGAA 1093
Qy
1026 GGCGGAGGCGCTTACCTTGGCCATGAGCTGAGTACCTTACCTAAGGAGCGGCTGAA 1085
Db
1094 GGAGTCTATCTCCAGGAGCAGCAGCTTCCAGCCGGAGTGTGAGGCGCCCTAGCC 1153
Qy
1086 GGAGTCTATCTCCAGGAGCAGCAGCTTCCAGCCGGAGTGTGAGGCGCCCTAGCC 1145
Db
1154 C 1154
Qy
1146 C 1146
Db

RESULT 12
US-09-949-016-4048
; Sequence 4048, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4048
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4048

Query Match 35.5%; Score 651.4; DB 3; Length 1701;
Best Local Similarity 74.4%; Pred. No. 2.8e-125;
Matches 819; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

Qy 33 GGCGGCGGGGCGGGGAGGCCCGTAGAACCGAGGGGGTCCGCCCGGGGGTCCCGGGGAG 92
Db 170 GGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 229

Qy 93 GTGGAGATGTTGAAGGGGCGAGCGGTTTCGAGCTGGGCGCGGGCTACACGCGAGTTGCGATAC 152
Db 230 CCGGAGATGTTCCCGCGCGCGAGGGTTCGAGCTGGGCGCGGGCTACACCAACCTCTCGTAC 289

Qy 153 ATCGGCGAGGCGCGTACGSCATGCTCGGCTATGACCACTGCGCAAGACTCGC 212
Db 290 ATCGGCGAGGCGCGCTACGSCATGCTGCTCTGCTATGATTAATGTCAACAAAGTTGCA 349

Qy 213 GTGGCCATCAAGAAGATCAAGCCCTTCGAACATCAGACCTTACTGCGAGCGCAGCTCCG 272
Db 350 GTAGCTATCAAGAAATCAAGCCCTTTGAGCACCAGACCTTACTGCGAGAGAACCTTGAG 409

Qy 273 GAGATCCAGATCTGCTGCGCTTCGCCATGAGAAATGTCATCGGCTCCGAGACATCTG 332
Db 410 GAGATAAAATCTTACTGCGCTTCAGACATGAGAAATCAATTTGGAATCAATGCAATTTAT 469

Qy 333 CGGCGCTCCACCTTGGAGCCATGAGAGTGTCTACATTTGTGCGAGGACCTGATGGAGACT 392
Db 470 CGAGCACCACCACTCAGGCAAAATGAAGATGTATATATAGTACAGGACCTCATGGAAACA 529

Qy 393 GACCTGTACAAGTTGCTGAAAAGCCAGCAGCTGAGCAATGACCAATATCTGCTACTTCTC 452
Db 530 GATCTTTACAAGCTCTTGAAGACACACCACTCAGCAATGACCAATATCTGCTATTTCTC 589

Qy 453 TACCAGATCTCGCGGCTCTCAAGTACATCCATCCGCGCAACGCTGCTCCACGAGATCTA 512
Db 590 TACCAGATCTCAGAGGGTTAAAATATATCCATTCAGCTAACGTTCTGCAACCGTGACCTC 649

Qy 513 AAGCCCTCCAACTGCTCATCAACACCACTGCGAGCTTAAGATTTGTGATTTTCGGCTG 572
Db 650 AAGCCCTCCAACTGCTGCTCAACACCACTGCTGATCTCAAGATCTGAGCTTTGGCTG 709

Qy 573 GCCGGATTGCCGATCTGAGCATGACCAACCGGCTTCTGAGGAGTATGTGGCTACG 632
Db 710 GCCGTTGTCAGATCCAGACCATGATCACAAGGGTTCCTGACAGAAATATGTGGCCACA 769

Qy 633 CGCTGTACCGGGCGCGGAGATCATGCTGAACTCCAAAGGGCTATACCAAGTCCATCGAC 692
Db 770 CGTTGGTACAGGGCTCCAGAAATTAATGTTGAAATTCAAAGGGCTACACCAAGTCCATTTGAT 829

Qy 693 ATCTGCTGTGGGCTGCAATCTGCTGAGATGCTCTCAACCGGCGCCATCTTCCCTGGC 752
Db 830 ATTTGCTGTAGGCTGCAATCTGCGCAGAAATGCTTCTAACAGGCGCCATCTTTCAGGG 889

Qy 753 AAGCCTTACCTGATCAGCTCAACCAATCTGCGGCTTCTGCGGCTCCCTCCATCCAGGAG 812
Db 890 AAGCAATATCTTGAACAGCTGAACCAATTTTGGTATCTTGGATCCCATCACAGAA 949

Qy 813 GACCTGAATTTGATCAATCAACATGAAGCCCGAAATCTACCTACAGTCTCTGCGCTCCAAG 872
Db 950 GACCTGAATTTGATCAATCAATGAAGCTAGGAACTATTTGCTTTCTCTCCACACAAA 1009

Qy 873 ACCAGTGGCTTGGGCGCAAGCTTTTCCCAAGTCTCAGACTCCAAAGCCCTTGACCTGCTG 932
Db

RESULT 14					
US-09-412-289-2					
; Sequence 2, Application US/09412289					
; Patent No. 6271210					
; GENERAL INFORMATION:					
; APPLICANT: Sivaraman, Vimala S.					
; APPLICANT: Wang, Hsien-Yu					
; APPLICANT: Malbon, Craig C.					
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES FOR MITOGEN-ACTIVATED					
; FILE REFERENCE: Seq. 1-4 (178-225 CIP II/CON)					
; CURRENT APPLICATION NUMBER: US/09/412,289					
; CURRENT FILING DATE: 1999-10-05					
; EARLIER APPLICATION NUMBER: 08/909,742					
; EARLIER FILING DATE: 1997-08-12					
; EARLIER APPLICATION NUMBER: 08/831,994					
; EARLIER FILING DATE: 1997-04-01					
; EARLIER APPLICATION NUMBER: 08/827,520					
; EARLIER FILING DATE: 1997-03-28					
; NUMBER OF SEQ ID NOS: 4					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 2					
; LENGTH: 1611					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
; FEATURE:					
; OTHER INFORMATION: Description of Artificial Sequence:blank					
US-09-412-289-2					
Qy	1	GAGAGTGGAGATGCGCGCGCGGGGTTCAGAGGGGGCGGGGGCGGGAGGCCCGTAGAA	60		
Dd	266	GAGGCGCGCGCCGCGAGCGTCGAGTCTCAGCGCGCGGAGGCGGGCGGCACGCA	325		
Qy	61	CCGAGGGGGTGCGGCCGGGGTCCCGGGGAGGTGGAGATGTTGAAGGGCAGCGGTTTCG	120		
Dd	326	ACATGGCGCGCGCGCGCGGGCGCGGGCCCGGAGATGGTCCGCGGGCAGGTGTTTCG	385		
Qy	121	ACGTGGGGCCGCGCTACACGCGATTGTCAGTACATCGCGAGGGCGCGTACGGCATGTGTA	180		
Dd	386	ACGTGGGGCCGCGCTACACCACCTCTCGTACATCGCGGAGGGCGCTACGGCATGGTGT	445		
Qy	181	GCTCGGCTATGACACCGTCGCGAGACTCGCGTGGCCATCAAGAAGATCAGCCCTTCG	240		
Dd	446	GCTCTGCTTATGATAATGTCAAAGTTCGAGTAGCTATCAAGAAAATCAGCCCCCTTG	505		
Qy	241	AACATCACACCTACTGCGCAGCGCAGCTCCGGGAGATCCAGATCTCTGTGGCTTCCGCC	300		
Dd	506	AGCACCGACCTACTGCGCAGAGAACCTTGAGGGAGATAAATCTTACTGCGCTTCAGAC	565		
Qy	301	ATGAGAATGTTCATCGGCATCCGAGACATTTCTCGGGCGTCCAACCTCGAAGCATGAGAG	360		
Dd	566	ATGAGAACATCATTTGGAATCAATGACATTTATTCGAGCACCAACCATCGAGCAATGAAG	625		
Qy	361	ATGCTACATTTGTGCGAGACCTGATGGAGACTGACAGTCTGTAAGTTGCTGAAAAGCCAGC	420		
Dd	626	ATGTATATATAGTACAGACCTCAT'GGAAACAAGATCTTTTACAAGCTCTTTGAAGACACAAC	685		
Qy	421	AGCTCAGCAATGACCATATCTGCTACTTCTCTACAGATCTCTGCGGGCGCTCAAGTACA	480		
Dd	686	ACCTCAGCAATGACCAATCTGCTATTTTCTTACAGATCTCTCAGAGGGTTAAATATA	745		
Qy	481	TCCACTCGGCCAACGTGCTCCAACGAGATCTTAAAGCCCTTCAAACCTCTGCTATCAACACA	540		

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 7, 2006, 12:49:28 ; Search time 1578.19 Seconds
(without alignments)
9625.515 Million cell updates/

Title: US-10-623-108-5
Perfect score: 1837
Sequence: 1 qaqqagtcgaqatggcgcgcg.....tagagatgtgtctatggctg 1837

Scoring table: IDENTITY_NUC

scoring cable: IDENTIFI₁ NOC
Gapop 10.0 ; Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Minimum DB seq length: 2000000000
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Full-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database :

```

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	1837	100.0	1837	8	US-10-623-108-5	Sequence 5, Appli	
2	1707	92.9	1777	8	US-10-623-108-7	Sequence 7, Appli	
3	1679	91.4	1869	5	US-10-098-841-114	Sequence 114, App	
4	1595	86.8	1726	8	US-10-623-108-3	Sequence 3, Appli	
5	1451	79.0	1654	8	US-10-623-108-1	Sequence 1, Appli	
6	1285.8	70.0	1400	10	US-10-060-756-1050	Sequence 1050, App	
7	1285.8	70.0	1400	10	US-11-060-756-5322	Sequence 5322, Ap	
8	1138.4	62.0	1896	5	US-10-072-036-38	Sequence 38, Appl	
9	1076	58.6	1214	3	US-09-925-297-323	Sequence 323, App	
10	1043.2	56.8	1147	7	US-10-114-270-109	Sequence 109, App	
11	958.6	52.2	1875	5	US-10-072-036-56	Sequence 56, Appl	
12	646.6	35.2	1611	5	US-10-171-311-126	Sequence 126, App	
13	646.6	35.2	1611	6	US-10-301-822-114	Sequence 114, App	
14	642	34.9	2791	6	US-10-305-720-1363	Sequence 1263, Ap	
15	642	34.9	2791	9	US-10-756-149-513	Sequence 513, App	
16	632.4	34.4	1074	9	US-10-451-291B-11	Sequence 11, Appl	
17	632.4	34.4	1815	5	US-10-072-036-58	Sequence 58, Appl	
18	632.4	34.4	1818	5	US-10-072-036-40	Sequence 40, Appl	
c 19	603.4	32.8	2111	4	US-09-925-065A-75282	Sequence 75282, A	
20	603.4	32.8	13808	3	US-09-764-877-3351	Sequence 3351, Ap	
21	603.4	32.8	13808	6	US-10-242-515-3351	Sequence 3351, Ap	
c 22	525.8	28.6	546	3	US-09-738-973-457	Sequence 457, App	
c 23	525.8	28.6	546	3	US-09-854-133-457	Sequence 457, App	

ALIGNMENTS

RESULT 1

REC'D
US-10-623-108-5

US-10-023-100-3
; Sequence 5, Application US/10623108

Publication No. US20050013817A1

; PUBLICATION NO: 0320
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: DAI, KEN-SHWO

APPLICANT: DAL, KEN-SHOW
TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS

; TITLE OF INVENTION: HUMAN
: FILE REFERENCE: U 014726-8

FILE REFERENCE: U 014726-8
CURRENT APPLICATION NUMBER: US/10/623,108

```

; ; CURRENT APPLICATION NUMBER: US/10/623,108
: : CURRENT FILING DATE: 2003-07-18

```

; CURRENT FILING DATE: 2003
 : NUMBER OF SEO ID NOS: 8

```

; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1

```

```
; SOFTWARE:
; SEQ ID NO 5
```

```
; SEQ ID NO 5
; LENGTH: 1837
```

```

; LENGTH: 1837
; TYPE: DNA

```

TYPE: DNA
ORGANISM: Homo sapiens

; ORGANISM: H

Query Match 100.0%; Score 1837; DB 8; Length 1837;

Query Match	100.0%;	Score	1337;
Best Local Similarity	100.0%;	Pred. No.	0;

BEST LOCAL SIMILARITY 100.0%; PRED. NO. 0;
Matches 1837: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAGAGTGGAGATGGCGGCGGCGCGCTCAGCGGGGCGGGGCGGGGAGCCCGGTAGAA	60
Db	1	GAGAGTGGAGATGGCGGCGGCGGCGCTCAGCGGGGCGGGGCGGGGAGCCCGGTAGAA	60
Qy	61	CCGAGGGGTCGGCCCGCGGGGTCGCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCGGTTCG	120
Db	61	CCGAGGGGTCGGCCCGCGGGGTCGCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCGGTTCG	120
Qy	121	ACGTGGGCGCGCCTACACGACGTTCAGGTACATCGCGAGGGGCGCTCAGGCATGTGCA	180
Db	121	ACGTGGGCGCGCCTACACGACGTTCAGGTACATCGCGAGGGGCGCTCAGGCATGTGCA	180
Qy	181	GCTCGGCCTATGACCACTGTCGCAAGACTCGCGTGGCCATCAAGAAGATCAGGCCCTTCG	240
Db	181	GCTCGGCCTATGACCACTGTCGCAAGACTCGCGTGGCCATCAAGAAGATCAGGCCCTTCG	240
Qy	241	AACATCAGACCTACTGCCAGCGACGCTCCGGGAGATCCAGATCCTCTCGCGTTCGCGCC	300
Db	241	AACATCAGACCTACTGCCAGCGACGCTCCGGGAGATCCAGATCCTCTCGCGTTCGCGCC	300
Qy	301	ATGAGAAATGTCATCGGCATCCGAGACATCTCGGGGCGTCCACCTCGGAAGCCATGAGAG	360
Db	301	ATGAGAAATGTCATCGGCATCCGAGACATCTCGGGGCGTCCACCTCGGAAGCCATGAGAG	360

Qy	361	ATGCTACATTGTGCGAGCACTGATGAGACTGACCTGTGTACAAAGTTGCTGAAAAGCCAGC	420
Db	361	ATGCTACATTGTGCGAGCACTGATGAGACTGACCTGTGTACAAAGTTGCTGAAAAGCCAGC	420
Qy	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACACAGATCTCCGCGGGGCTCTCAAGTACA	480
Db	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACACAGATCTCCGCGGGGCTCTCAAGTACA	480
Qy	481	TCCACTCCGCCAAACGTGTCTCACCGAGATCTAAAGCCCTCAACCTGCTCATCAACACCA	540
Db	481	TCCACTCCGCCAAACGTGTCTCACCGAGATCTAAAGCCCTCAACCTGCTCATCAACACCA	540
Qy	541	CCTCGGACCTTAAGATTGTGATTTGGCGCTGGCCCGGATTCGGAGTCCTGAGCATGACC	600
Db	541	CCTCGGACCTTAAGATTGTGATTTGGCGCTGGCCCGGATTCGGAGTCCTGAGCATGACC	600
Qy	601	ACACCGGCTTCTCACCGGAGTATGTGCTACGCGCTGTACCGGGCCCCAGAGATCATGC	660
Db	601	ACACCGGCTTCTCACCGGAGTATGTGCTACGCGCTGTACCGGGCCCCAGAGATCATGC	660
Qy	661	TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGTGGGCTGATCTTGGCTG	720
Db	661	TGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGTGGGCTGATCTTGGCTG	720
Qy	721	AGATGCTCTTAAACCGGCCCATTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACCA	780
Db	721	AGATGCTCTTAAACCGGCCCATTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACCA	780
Qy	781	TTCTGGGCATCTGGGCTCCCATCCAGAGGACCTGAATTGTATCATCAACATGAAGG	840
Db	781	TTCTGGGCATCTGGGCTCCCATCCAGAGGACCTGAATTGTATCATCAACATGAAGG	840
Qy	841	CCCGAAACTACTACAGTCTCTGCCCTCCAAGACCAAGGTGGCTTGGGCCCAAGCTTTTCC	900
Db	841	CCCGAAACTACTACAGTCTCTGCCCTCCAAGACCAAGGTGGCTTGGGCCCAAGCTTTTCC	900
Qy	901	CCAAAGTCAGACTCCAAAGCCCTTGACCTGTCTGGAACCGGATGTTAACTTTAAACCCCAATA	960
Db	901	CCAAAGTCAGACTCCAAAGCCCTTGACCTGTCTGGAACCGGATGTTAACTTTAAACCCCAATA	960
Qy	961	AACCGATCAGTGGAGGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA	1020
Db	961	AACCGATCAGTGGAGGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA	1020
Qy	1021	CGGATGAGCCAGTGGCGAGGAGCCCTTACCTTCGCATCGAGCTGGATGACCTACTTA	1080
Db	1021	CGGATGAGCCAGTGGCGAGGAGCCCTTACCTTCGCATCGAGCTGGATGACCTACTTA	1080
Qy	1081	AGGAGCGGCTGAAGGAGTCACTTTCAGGAGACAGCACGCTTCACGCCGAGTGTCTGG	1140
Db	1081	AGGAGCGGCTGAAGGAGTCACTTTCAGGAGACAGCACGCTTCACGCCGAGTGTCTGG	1140
Qy	1141	AGGCCCTCTAGCCAGACAGACATCTGTGCACCTTGGGGCTTGGAAACAGAACTGGCAAG	1200
Db	1141	AGGCCCTCTAGCCAGACAGACATCTGTGCACCTTGGGGCTTGGAAACAGAACTGGCAAG	1200
Qy	1201	AGGCAAGAGGTCACTGAGGGGCTGTGTACCCAGGACCTGCTCTCTGCCCTCTCTCC	1260
Db	1201	AGGCAAGAGGTCACTGAGGGGCTGTGTACCCAGGACCTGCTCTCTGCCCTCTCTCC	1260
Qy	1261	CGCCAGACTGTTAGAAAATGGACACTGTGCCACCGCGACTTTGGCAGGCCACAGCCGGG	1320
Db	1261	CGCCAGACTGTTAGAAAATGGACACTGTGCCACCGCGACTTTGGCAGGCCACAGCCGGG	1320
Qy	1321	GTGAGCATGGGCTGGGCCACTCTCTCTTTTGTGAGGGCTTCAGCTTCAGGCAAGGCCA	1380
Db	1321	GTGAGCATGGGCTGGGCCACTCTCTCTTTTGTGAGGGCTTCAGCTTCAGGCAAGGCCA	1380
Qy	1381	AGGCTCTTCTCTCCCAACCGCCCTCCCAAGGGGCTTGGGAGCTCAGGTGGCCCCAGT	1440
Db	1381	AGGCTCTTCTCTCCCAACCGCCCTCCCAAGGGGCTTGGGAGCTCAGGTGGCCCCAGT	1440
Qy	1441	TCAATCTCCCGCTGTGTGTGGCGCCCTTACTTTCCCAAGCTGTCAGTCTCTGGCAGT	1500

RESULT 2

US-10-623-108-7

US-10-623-108-7
: Sequence 7. Application US/10623108; sequence 7, Application US/1062
: Publication No. US20050013817A1

Publication No: US20060136401A1

```

; GENERAL INFORMATION:
; APPLICANT: DAT. KEN-SHOW

```

APPLICANT: DAI, KEN-SHOU
TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS

; TITLE OF INVENTION: HUMAN
: FILE REFERENCE: U 014726-8

FILE REFERENCE: 0014726-8
CURRENT APPLICATION NUMBER: US/10/623.108

; CURRENT APPLICATION NUMBER: US/18-07-18
 : CURRENT FILING DATE: 2003-07-18

```

; CURRENT FILING DATE: 20
; NUMBER OF SEC ID NOS: 8

```

; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1.1; SOFTWARE: L
; SEO ID NO 7; SEQ ID NO: 1
; LENGTH: 1777

TYPE: DNA

LIFE: DNA
 ORGANISM: Homo sapiens

US-10-623-108-7

Query Match 92.9%; Score 1707; DB 8; Length 1777;

Query Match 52.5%; Pred. No. 0;
Best Local Similarity 96.7%; Pred. No. 0;

BEST LOCAL SIMILARITY 50.7%; P-ED: NO: 0;
 Matches 1777: Conservative 0; Mismatches 0; Indels 60; Gaps 1;

1 GAGGAGTGGAGATGGCGCGCGCGCGCTCAGGGGGCGGGGAGCCCGTAGAA 60

1 CAGCGCTGCGACATGGCCGCCGCCGCCTCAGGGGGCGGGGAGCCCCGTAGAA 60

12

[illegible]

PROPERTY OF

DD T ZT ACATACCCCTCACCATTGTCAGTTGGCGGTGC

QY GCCTCGGCTTAATGACCACTGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAAGCCCTTCC
I8I

Db 181 GCTCGGCCTATGACCAACGTGCGCAAGACTTCGCCGTGGCCATCAAGAAAGATCAGCCCCCTCG 24

QY 241 AACATCAGACCTACTGCCAGGCACGCTCCGGGAGATCCAGATCCCTGCTGGCGTTCCGCC 30

Db 241 AACATCAGACCTACTGCCAGCGCACGCTCCGGAGATCCAGATCCTGCTGCGCTTCCGCC 30

QV 301 ATGAGAAATGTCATCGGCATCCGAGACATTTCTGCGGGCGTCCACCCCTGGAAGCCATGAGAG 36

```
Qy 361 ATGTCTACATTTGTGACGAGCCTGATGGAGACTGACCTGTACAAGTTGCTGAAAGCCAGC 420
Db 361 ATGTCTACATTTGTGACGAGCCTGATGGAGACTGACCTGTACAAGTTGCTGAAAGCCAGC 420
Qy 421 AGCTGACAAATGACCATATCTGCTACTCTCTACAGATCTCGCGGGCTCAAGTACA 480
Db 421 AGCTGACAAATGACCATATCTGCTACTCTCTACAGATCTCGCGGGCTCAAGTACA 480
Qy 481 TCCACTCCGCAACGCTCTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
Db 481 TCCACTCCGCAACGCTCTCCACGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
Qy 541 CTTGCGACCTTAAGATTGTGAATTTGGCTTGGCCCGGATTTGGCGATCTCTGAGCATGACC 600
Db 541 CTTGCGACCTTAAGATTGTGAATTTGGCTTGGCCCGGATTTGGCGATCTCTGAGCATGACC 600
Qy 601 ACACCGGCTTCTGACCGAGTATGTGGCTAGCGCTGGTACCGGCCCCAGAGATCATGC 660
Db 601 ACACCGGCTTCTGACCGAGTATGTGGCTAGCGCTGGTACCGGCCCCAGAGATCATGC 660
Qy 661 TGAATCTCAAGGGCTATACCAAGTCCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 720
Db 661 TGAATCTCAAGGGCTATACCAAGTCCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTG 720
Qy 721 AGATGCTCTCTAAACCGGCCCATTTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 780
Db 721 AGATGCTCTCTAAACCGGCCCATTTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA 780
Qy 781 TTCTGGGATCTCTGGGCTCCCATCCAGAGGAGCTGAAATGTATCATCAACATGAAG 840
Db 781 TTCTGGGATCTCTGGGCTCCCATCCAGAGGAGCTGAAATGTATCATCAACATGAAG 840
Qy 841 CCCGAACCTACCTACAGTCTCTGCTCCCTCCAGACCAAGGTGGCTTGGCCAAAGCTTTTCC 900
Db 841 CCCGAACCTACCTACAGTCTCTGCTCCCTCCAGACCAAGGTGGCTTGGCCAAAGCTTTTCC 900
Qy 901 CCAAGTCAGATCTCAAGAGCCCTTGAACCTGTCTGACCGGATGTAACTTTAACCCCAATA 960
Db 901 CCAAGTCAGATCTCAAGAGCCCTTGAACCTGTCTGACCGGATGTAACTTTAACCCCAATA 960
Qy 961 AACCGATCAGATGGAGAGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA 1020
Db 961 AACCGATCAGATGGAGAGAGCGCTGGCTACCCCTACCTGGAGCAGTACTATGACCCGA 1020
Qy 1021 CGGATGAGCAGTGGCGGAGAGCGCTTACCTTCCGCAATGAGCTGATGACCTACTTA 1080
Db 1021 CGGATGAGCAGTGGCGGAGAGCGCTTACCTTCCGCAATGAGCTGATGACCTACTTA 1080
Qy 1081 AGGAGCGCTGAAGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1140
Db 1081 AGGAGCGCTGAAGAGCTCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGTGG 1140
Qy 1141 AGGCCCCCTAGCCAGACAGATCTCTGCACTTCCAGGAGACAGCAGCTTCCAGGAG 1200
Db 1141 AGGCCCCCTAGCCAGACAGATCTCTGCACTTCCAGGAGACAGCAGCTTCCAGGAG 1200
Qy 1201 AGGCAAGAGTCTAGTGGGCTCTGTACCCAGGAGCTGCTCTGCTGCTGCTGCTCTCC 1260
Db 1201 AGGCAAGAGTCTAGTGGGCTCTGTACCCAGGAGCTGCTCTGCTGCTGCTGCTCTCC 1260
Qy 1261 CGCCAGACTGTTAGAAATGACACTGTGCGCCAGCCCGGAGCTTGGCAGCCCGGCGGG 1320
Db 1261 CGCCAGACTGTTAGAAATGACACTGTGCGCCAGCCCGGAGCTTGGCAGCCCGGCGGG 1320
Qy 1321 GTGAGCATGGGCTTGGCCACTCTCTCTCTTGTGAGGCTTCCAGCTTTCAGGAGGCCA 1380
Db 1321 GTGAGCATGGGCTTGGCCACTCTCTCTCTTGTGAGGCTTCCAGCTTTCAGGAGGCCA 1380
Qy 1381 AGGCTTCTCTCCACCGGCTTCCCAAGGGCTTCCGAGCTCAGGTGGGCCCCAGT 1440
Db 1381 AGGCTTCTCTCCACCGGCTTCCCAAGGGCTTCCGAGCTCAGGTGGGCCCCAGT 1440
```

```
Qy 1441 TCAATCTCCGCTGCTGCTGCGCCCTTACCTTCCAGCGTCCAGTCTCTGSCAGT 1500
Db 1441 TCAATCTCCGCTGCTGCTGCGCCCTTACCTTCCAGCGTCCAGTCTCTGSCAGT 1440
Qy 1501 TCTGGAATGGAAGGGTTCTGGCTGCGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGG 1560
Db 1441 TCTGGAATGGAAGGGTTCTGGCTGCGCCCAACCTGCTGAAGGGCAGAGGTGGAGGGTGGG 1500
Qy 1561 GGCGCTGAGTGAAGGATCTCAGGGCCATGCTGCGCCCTCATCTCATTTCAAACCCACCT 1620
Db 1501 GGCGCTGAGTGAAGGATCTCAGGGCCATGCTGCGCCCTCATCTCATTTCAAACCCACCT 1560
Qy 1621 AGTTTCCCTGAAGGAACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCCG 1680
Db 1561 AGTTTCCCTGAAGGAACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCCG 1620
Qy 1681 GGCGGAATCCCTTCCCTGTCANAGCTGTCACTTTCGGGTGCCCTCGCTGCTTCTGTGTG 1740
Db 1621 GGCGGAATCCCTTCCCTGTCANAGCTGTCACTTTCGGGTGCCCTCGCTGCTTCTGTGTG 1680
Qy 1741 GTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCCCGGCCCTTGGCACCTTCCCTGACC 1800
Db 1681 GTGAGCAGAGTGGAGCTGGGGGGCTGGAGAGCCCGGCCCTTGGCACCTTCCCTGACC 1740
Qy 1801 CGTCTAATATATATATATATAGATGTGTCTATGCTG 1837
Db 1741 CGTCTAATATATATATATATAGATGTGTCTATGCTG 1777
```

RESULT 3
US-10-098-841-114
; Sequence 114, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 114
; LENGTH: 1869
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)..(1212)
US-10-098-841-114

Query Match 91.4%; Score 1679.8; DB 5; Length 1869;

Best Local Similarity 96.6%; Pred. No. 0; Matches 1782; Conservative 0; Mismatches 2; Indels 61; Gaps 4;									
QY	3	GGAGTGGAGATGCGCGCGCGCGGTCTAGGGGGGCGGGGGCGGGAGCCCGTGTAGAACC	62						
Db	64	GGAGTGGAGATGCGCGCGCGCGGTCTAGGGGGGCGGGGGCGGGAGCCCGTGTAGAACC	123						
QY	63	GAGGGGTGGCCCGGGGGTCCCGGGGAGGTGGAGATGTTGAAGGGGAGCCGTTTCGAC	122						
Db	124	GAGGGGTGGCCCGGGGGTCCCGGGGAGGTGGAGATGTTGAAGGGGAGCCGTTTCGAC	183						
QY	123	GTGGGCGCGCGCTACACGAGTGTGAGTATCATCGGCGAGGGCGGTACGGCATGTGTGACG	182						
Db	184	GTGGGCGCGCGCTACACGAGTGTGAGTATCATCGGCGAGGGCGGTACGGCATGTGTGACG	243						
QY	183	TCGGCCTATGACCACTGTGGCGAAGACTCGGTGGCCATCAAGAGATCAGCCCTTCGAA	242						
Db	244	TCGGCCTATGACCACTGTGGCGAAGACTCGGTGGCCATCAAGAGATCAGCCCTTCGAA	303						
QY	243	CATCAGACCTACTGCGCAGCGACTTCGCGGGCGTCCACCTGGAAGCCATGAGAGAT	362						
Db	304	CATCAGACCTACTGCGCAGCGACTTCGCGGGAGATCCAGATCTGTGCGCTTCGCGCAT	363						
QY	303	GAGATGTTCATCGGCATCCGAGACATTCGCGGGCGTCCACCTGGAAGCCATGAGAGAT	423						
Db	364	GAGATGTTCATCGGCATCCGAGACATTCGCGGGCGTCCACCTGGAAGCCATGAGAGAT	423						
QY	363	GTCTAATTTGTGAGGACCTGATGAGAGTGAAGTCTGTAAGTTGTGAAAGCCAGGAG	422						
Db	424	GTCTAATTTGTGAGGACCTGATGAGAGTGAAGTCTGTAAGTTGTGAAAGCCAGGAG	483						
QY	423	CTGAGCAATGACCATCTGCTACTTCTCTACACAGATCTCGGGGGCTCAAGTATCATC	482						
Db	484	CTGAGCAATGACCATCTGCTACTTCTCTACACAGATCTCGGGGGCTCAAGTATCATC	543						
QY	483	CACCTCCGCCAACTGTCTCAACGAGATCTAAAGCCCTCCAACTGTCTATCAACACACC	542						
Db	544	CACCTCCGCCAACTGTCTCAACGAGATCTAAAGCCCTCCAACTGTCTATCAACACACC	603						
QY	543	TGCGAATTAAGATTTGTGATTTTCGGCTGGCGCCGGAATTCGCGATCTGAGCATGACCAC	602						
Db	604	TGCGAATTAAGATTTGTGATTTTCGGCTGGCGCCGGAATTCGCGATCTGAGCATGACCAC	663						
QY	603	ACCGGCTTCTGACGAGATGTGGCTACGCGTGTACCGGCGCCAGAGATCATGCTG	662						
Db	664	ACCGGCTTCTGACGAGATGTGGCTACGCGTGTACCGGCGCCAGAGATCATGCTG	723						
QY	663	AACCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTTCTGGCTGAG	722						
Db	724	AACCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATTTCTGGCTGAG	783						
QY	723	ATGCTCTTAAACGGGCCATCTTCCCTTGGCAAGCACTACCTGGATCAGCTCAACCAATT	782						
Db	784	ATGCTCTTAAACGGGCCATCTTCCCTTGGCAAGCACTACCTGGATCAGCTCAACCAATT	843						
QY	783	CTGGGATCTGGGGTCCCATCCAGAGGACCTGAATTTGATATCAATCAACAGAGCC	842						
Db	844	CTGGGATCTGGGGTCCCATCCAGAGGACCTGAATTTGATATCAATCAACAGAGCC	903						
QY	843	CGAAACTACCTACAGTCTCTGCGCTCCAAAGCAAGGTGGCTTTGGGCCAAGCTTTTCCCC	902						
Db	904	CGAAACTACCTACAGTCTCTGCGCTCCAAAGCAAGGTGGCTTTGGGCCAAGCTTTTCCCC	963						
QY	903	AAGTCAGATCTCAAAGCCCTTGACTGTGGACCGGATGTAACTTTAACCCCAATAA	962						
Db	964	AAGTCAGATCTCAAAGCCCTTGACTGTGGACCGGATGTAACTTTAACCCCAATAA	1023						
QY	963	CGGATCAGAGTGAGGAGCGCTGCGCTCACCCCTACTCGGAGCAGTATGACCCGACG	1022						
Db	1024	CGGATCAGAGTGAGGAGCGCTGCGCTCACCCCTACTCGGAGCAGTATGACCCGACG	1083						
QY	1023	GATGAGCAGTGGCCGAGAGCCCTTACCTTCGCCATGAGCTGGATGACCTACCTAAG	1082						

RESULT 4

US-10-623-108-3
; Sequence 3, Application US/10623108
; Publication NO. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHMO
; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1726
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-108-3

Db	1084	GATGAGCCAGTGGCGAGGAGCCCTTACCTTCGCCATGAGCTGGATGACCTACCTAAG	1143
QY	1083	GAGCGGCTGAGGAGCTCATCTTCCAGAGAGACAGACGCTTCCAGCCGGAGTGTCTGGAG	1142
Db	1144	GAGCGGCTGAGGAGCTCATCTTCCAGAGAGACAGACGCTTCCAGCCGGAGTGTCTGGAG	1203
QY	1143	GGCCCCCTAG-CCGACAGACAGACATCTCTGACCCCTGGGCGCTTGGAAACAGAACTGGCAAGA	1201
Db	1204	GGCCCCCTAGCCGACAGACAGACATCTCTGACCCCTGGGCGCT-----	1244
QY	1202	GGCAAGAGGTACTGAGGGCCCTCTGTCAACCCAGGACCTGCGCTCTGCTGCGCCCTCTCCC	1261
Db	1245	-----GGACCTGCGCTCTGCGCCCTCTCCC	1272
QY	1262	GCCAGACTGTTAGAAAATGGACACATGTGCCAGCCGAGACCTTGGCAGCCAGGCGGGG	1321
Db	1273	GCCAGACTGTTAGAAAATGGACACATGTGCCAGCCGAGACCTTGGCAGCCAGGCGGGG	1332
QY	1322	TGGAGCATGGGCTTGGCCACCTCTCTCTCTTGTCTGAGGCTCTCAGCTTTCAGGC-----A	1375
Db	1333	TGGAGCATGGGCTTGGCCACCTCTCTCTCTTGTCTGAGGCTCTCAGCTTTCAGGCAGGCCAA	1392
QY	1376	GGCCAAAGGCTTCTCTCTCCACCCGCTCTCCGAGGGGCTCTCGGAGCTCAGGTGGCC	1435
Db	1393	GGCCAAAGGCTTCTCTCTCCACCCGCTCTCCGAGGGGCTCTCGGAGCTCAGGTGGCC	1452
QY	1436	CCAGTTCATCTCCC---GCTGCTGCTGTCGGCCCTTACCTTCCCGAGCTCCAGTCT	1492
Db	1453	CCAGTTCATCTCCCCTGCTGCTGTCGGCCCTTACCTTCCCGAGCTCCAGTCT	1512
QY	1493	CTGGCAGTCTTGGAAATGGAAAGGTTCTGGCTGCCCCCAACCTGCTGAAAGGCGAGAGTGA	1552
Db	1513	CTGGCAGTCTTGGAAATGGAAAGGTTCTGGCTGCCCCCAACCTGCTGAAAGGCGAGAGTGA	1572
QY	1553	GGGTGGGGGGGCTGAGTAGGAGCTCAGGGCCATGCGCTGCCCCCTCATCTCATTTCAAC	1612
Db	1573	GGGTGGGGGGGCTGAGTAGGAGCTCAGGGCCATGCGCTGCCCCCTCATCTCATTTCAAC	1632
QY	1613	CCACCCCTAGTTTCCCTGAAAGGAACATTCCTAGTCTCAAGGGCTAGCATCCCTGAGGAG	1672
Db	1633	CCACCCCTAGTTTCCCTGAAAGGAACATTCCTAGTCTCAAGGGCTAGCATCCCTGAGGAG	1692
QY	1673	CCAGGCGGGGCGAAATCCCTCTGTCAAAAGCTGTCACTTCGCGTGGCCCTCGCTGCTTC	1732
Db	1693	CCAGGCGGGGCGAAATCCCTCTGTCAAAAGCTGTCACTTCGCGTGGCCCTCGCTGCTTC	1752
QY	1733	TGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGGCTGGAGAGCCCGGCGCCCTTGCACCT	1792
Db	1753	TGTGTGTGTGAGCAGAAAGTGGAGCTGGGGGGCTGGAGAGCCCGGCGCCCTTGCACCT	1812
QY	1793	CCCTGACCGCTCTAATATATATATATATAGAGATGTGTCTATGGCTG	1837
Db	1813	CCCTGACCGCTCTAATATATATATATATAGAGATGTGTCTATGGCTG	1857

Query Match		86.8%; Score 1595; DB 8; Length 1726;
Best Local Similarity		94.0%; Pred. No. 0;
Matches 1726; Conservative		0; Mismatches 0; Indels 111; Gaps 2;
Qy	1	GAGGAGTGGAGATGGCGCGCGCGCTCAGGGGGGCGGGGCGGGGAGCCCGTAGAA 60
Db	1	GAGGAGTGGAGATGGCGCGCGCGCGCTCAGGGGGGCGGGGCGGGGAGCCCGTAGAA 60
Qy	61	CCGAGGGGGTCCGCGCGCGCGCGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTGG 120
Db	61	CCGAGGGGGTCCGCGCGCGCGCGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCCGTTGG 120
Qy	121	ACGTGGCGCGCGCTACACGAGTTGACATACATCGCGAGGGGCGGTACGGCATGTGTCA 180
Db	121	ACGTGGCGCGCGCTACACGAGTTGACATACATCGCGAGGGGCGGTACGGCATGTGTCA 180
Qy	181	GCTCGGCGTATGACCGAGTGGCGGAGACTCGCGTGGCCATCAAGAGATCAGCCCTTCG 240
Db	181	GCTCGGCGTATGACCGAGTGGCGGAGACTCGCGTGGCCATCAAGAGATCAGCCCTTCG 240
Qy	241	AACATCAGAGCTTACTGCGAGCGACGCTCCGGGAGATCCAGATCCTGCTCGCTTCGGCC 300
Db	241	AACATCAGAGCTTACTGCGAGCGACGCTCCGGGAGATCCAGATCCTGCTCGCTTCGGCC 300
Qy	301	ATGAGAAATGTCGGCATCCGAGACATTCCTGGGGGCGTCCACCTGGAAAGCCATGAGAG 360
Db	301	ATGAGAAATGTCGGCATCCGAGACATTCCTGGGGGCGTCCACCTGGAAAGCCATGAGAG 360
Qy	361	ATGCTACATTTGTCAGGACTGATGAGAGTGAACCTGTACAAAGTTGCTGAAAAGCCAGC 420
Db	361	ATGCTACATTTGTCAGGACTGATGAGAGTGAACCTGTACAAAGTTGCTGAAAAGCCAGC 420
Qy	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACAGATCCTGCGGGGCTCAAGTACA 480
Db	421	AGCTGAGCAATGACCATATCTGCTACTTCTCTACAGATCCTGCGGGGCTCAAGTACA 480
Qy	481	TCCACTCCGCGCAACGCTGCTCCAGAGATTAAGCCCTCCAACTGCTCATCAACACCA 540
Db	481	TCCACTCCGCGCAACGCTGCTCCAGAGATTAAGCCCTCCAACTGCTCATCAACACCA 540
Qy	541	CCTCGGACCTTAAAGTTGATTTGCGCTGCGCGCGCGGATTCGGATCCCTGAGCATGACC 600
Db	541	CCTCGGACCTTAAAGTTGATTTGCGCTGCGCGCGCGGATTCGGATCCCTGAGCATGACC 600
Qy	601	ACACCGCTTCTGACGAGTATGCTACGCGCTGCTACCGGGGCGCGAGATCATGC 660
Db	601	ACACCGCTTCTGACGAGTATGCTACGCGCTGCTACCGGGGCGCGAGATCATGC 660
Qy	661	TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTGGCTG 720
Db	661	TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTGGCTG 720
Qy	721	AGATGCTCTTAACCGGCGCATCTTCCCTGGCAGCACTACCTGGATCAGCTCAACACA 780
Db	721	AGATGCTCTTAACCGGCGCATCTTCCCTGGCAGCACTACCTGGATCAGCTCAACACA 780
Qy	781	TTCTGGGATCCTGGGCTCCCATCCAGAGGACCTGAATTTGATCATCAATGAAGG 840
Db	781	TTCTGGGATCCTGGGCTCCCATCCAGAGGACCTGAATTTGATCATCAATGAAGG 840
Qy	841	CCCGAAACTACCTACAGTCTCTGCGCTCCAAAGCAAGGTGGCTTGGGCCAAGCTTTTC 900
Db	841	CCCGAAACTACCTACAGTCTCTGCGCTCCAAAGCAAGGTGGCTTGGGCCAAGCTTTTC 900
Qy	901	CCAAGTGAGCTCAAGCCCTTGAACCTGCGGAGCGGATGTTAACTTTAACCCCAATA 960
Db	901	CCAAGTGAGCTCAAGCCCTTGAACCTGCGGAGCGGATGTTAACTTTAACCCCAATA 960
Qy	961	AACGGATCAGTGGAGGAGCGGTGGCTCACCCCTACCTGAGCAGTACTATGACCCGA 1020
Db	961	AACGGATCA----- 969
Qy	1021	CGGATGAGCCAGTGGCGGAGGAGCCCTTACCTTCCGCGATGGAGCTGGATGACCTACTA 1080

Db	970	-----CAGTGGCGGAGGAGCCCTTACCTTCCGATGGAGCTGGATGACCTACTACTA 1020
Qy	1081	AGGAGCGGCTGAAGAGTCTATCTTCCAGGAGACAGACGCTTCCAGCCCGAGTGTGG 1140
Db	1021	AGGAGCGGCTGAAGAGTCTATCTTCCAGGAGACAGACGCTTCCAGCCCGAGTGTGG 1080
Qy	1141	AGGCCCTTAGCCAGACAGACATCTCTGACCCCTGGGGCTGGAAACAGACTGGCNAAG 1200
Db	1081	AGGCCCTTAGCCAGACAGACATCTCTGACCCCTGGGGCT----- 1122
Qy	1201	AGGCAAGAGTCACTGAGGGCTCTCTCAACCAGGAGCTTGCCTCTCTGCTGCTCTCC 1260
Db	1123	-----GACCTTGCCTCTCTGCTGCTCTCTCTCC 1149
Qy	1261	CGCCAGACTGTTAGAAAATGGACATGTGTGCCAGCCCGGACCTTGGCAGCCAGGCGGG 1320
Db	1150	CGCCAGACTGTTAGAAAATGGACATGTGTGCCAGCCCGGACCTTGGCAGCCAGGCGGG 1209
Qy	1321	GTGGAGATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTCCAGCTTCAGGACAGGCA 1380
Db	1210	GTGGAGATGGGCTGGCCACCTCTCTCTTGTGTGAGGCTCCAGCTTCAGGACAGGCA 1269
Qy	1381	AGGCTTCTCTCCACCCAGCCGCTCCCAAGGGGCTCGGGAGCTCAGGTGCCCCAGT 1440
Db	1270	AGGCTTCTCTCTCCACCCGCTCCCAAGGGGCTCGGGAGCTCAGGTGCCCCAGT 1329
Qy	1441	TCAATCTCCGCTGCTGCTGCTGCGCCCTTACCTTCCAGCGCTCCAGCTCTTGGCAGT 1500
Db	1330	TCAATCTCCGCTGCTGCTGCTGCGCCCTTACCTTCCAGCGCTCCAGCTCTTGGCAGT 1389
Qy	1501	TCTGGAATGGAAGGTTCTGGCTGCCCCCAACCTGCTGAAGGCGAGAGTGGAGGTGGG 1560
Db	1390	TCTGGAATGGAAGGTTCTGGCTGCCCCCAACCTGCTGAAGGCGAGAGTGGAGGTGGG 1449
Qy	1561	GGCGCTGAGTGAAGGCTCAGGGGCTGCTGCCCCCTCATCTCATTCATTAACCCACCT 1620
Db	1450	GGCGCTGAGTGAAGGCTCAGGGGCTGCTGCCCCCTCATCTCATTCATTAACCCACCT 1509
Qy	1621	AGTTTCCCTGGAAGAACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGAGCGAGCGG 1680
Db	1510	AGTTTCCCTGGAAGAACATTTCTTAGTCTCAAGGGCTAGCATCCCTGAGAGCGAGCGG 1569
Qy	1681	GGCGGAATCCCTTCCCTGTCAAAAGCTGTCACTTGGGTGCGCTGCTGCTGCTGTGTG 1740
Db	1570	GGCGGAATCCCTTCCCTGTCAAAAGCTGTCACTTGGGTGCGCTGCTGCTGCTGTGTG 1629
Qy	1741	GTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCGCGGCCCTGCGACCTCCCTGACC 1800
Db	1630	GTGAGCAGAGTGGAGCTGGGGGCGTGGAGAGCCGCGGCCCTGCGACCTCCCTGACC 1689
Qy	1801	CGTCTAATATATAAATATAGAGATGTCTTATGGCTG 1837
Db	1690	CGTCTAATATATAAATATAGAGATGTCTTATGGCTG 1726

RESULT 5

US-10-623-108-1
; Sequence 1, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAL, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-108-1

Query Match		79.0%;	Score 1451;	DB 8;	Length 1654;
Best Local Similarity		90.0%;	Pred. No. 0;		
Matches 1654; Conservative		0;	Mismatches	0;	Indels 183; Gaps 2;
QY	1	GAGAGTGGAGATGGCGGGCGCGGCTCAGGGGGCGGGGCGGGAGCCCCGTAGAA	60		
DB	1	GAGAGTGGAGATGGCGGGCGCGGCTCAGGGGGCGGGGCGGGAGCCCCGTAGAA	60		
QY	61	CGAGGGGTGCGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCGAGCCGTTGG	120		
DB	61	CGAGGGGTGCGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCGAGCCGTTGG	120		
QY	121	AGTGGGCGCGCTACACGAGTTCAGTACATCGGCGAGGCGGCTACGGCATGGTCA	180		
DB	121	ACGTGGGCGCGCTACACGAGTTCAGTACATCGGCGAGGCGGCTACGGCATGGTCA	180		
QY	181	GCTCGGCTATGACACGTCGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCCCTCG	240		
DB	181	GCTCGGCTATGACACGTCGCGCAAGACTCGCGTGGCCATCAAGAAAGATCAGCCCCCTCG	240		
QY	241	AACATCAGACCTACTGCGAGCGCACGCTCGGGAGATCCAGATCCCTGCGCTTCGCGC	300		
DB	241	AACATCAGACCTACTGCGAGCGCACGCTCGGGAGATCCAGATCCCTGCGCTTCGCGC	300		
QY	301	ATGAGAATGTCATCGGCATCCGAGACATTTCTGGGGCGTCCACCTGGGAAGCCATGAGAG	360		
DB	301	ATGAGAATGTCATCGGCATCCGAGACATTTCTGGGGCGTCCACCTGGGAAGCCATGAGAG	360		
QY	361	ATGTCTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAGCCAGC	420		
DB	361	ATGTCTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAGCCAGC	420		
QY	421	AGCTGAGCAATGACCAATATCTGCTACTTCTCTACCAAGATCCCTGCGGGGCTCAAGTACA	480		
DB	421	AGCTGAGCAATGACCAATATCTGCTACTTCTCTACCAAGATCCCTGCGGGGCTCAAGTACA	480		
QY	481	TCCACCTCCGCAACGTCCTCCACGAGATCTAAAGCCCTCCACCTGCTCATCAACACCA	540		
DB	481	TCCACCTCCGCAACGTCCTCCACGAGATCTAAAGCCCTCCACCTGCTCATCAACACCA	540		
QY	541	CCTGCGACCTTAAGATTTGTGATTTTCGGCTGCGCCGGATGCGCATCTCGAGCATGACC	600		
DB	541	CCTGCGACCTTAAGATTTGTGATTTTCGGCTGCGCCGGATGCGCATCTCGAGCATGACC	600		
QY	601	ACACCGGCTTCTGACGGAGTATGCGCTACGCGCTGGTACCGGGCCCCAGAGATCATGC	660		
DB	601	ACACCGGCTTCTGACGGAGTATGCGCTACGCGCTGGTACCGGGCCCCAGAGATCATGC	660		
QY	661	TGAATCCAGGGCTATACCAAGTCCATCGACATCTGCTGCGGCTGCATTTCTGGCTG	720		
DB	661	TGAATCCAGGGCTATACCAAGTCCATCGACATCTGCTGCGGCTGCATTTCTGGCTG	720		
QY	721	AGATGCTCTTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA	780		
DB	721	AGATGCTCTTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATCAGCTCAACACA	780		
QY	781	TTCTGGGCTATCCGGGCTCCCCATCCAGGAGGACCTGAAATGTATCATCAATGAAGG	840		
DB	781	TTCTG-----	785		
QY	841	CCGAAATACCTACAGTCTCTGCCCTCCAAAGACCAAGTGGCTTGGGCCAAGCTTTTCC	900		
DB	786	-----	785		
QY	901	CCAAGTCAGATCCAAAGCCCTTACCTGTGACCCGGATGTTAACTTTAACCCCAATA	960		
DB	786	-----GCCCTTGACCTGTGACCCGGATGTTAACTTTAACCCCAATA	828		
QY	961	AACGGATCAGTGGAGGAAGCCCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA	1020		
DB	829	AACGGATCAGTGGAGGAAGCCCTGGCTCACCCCTACCTGGAGCAGTACTATGACCCGA	888		

QY	1021	CGGATGAGCCAGTGGCCGAGGAGCCCTTACCTTCGCATGGAGTGGATGACCTACCTTA	1080
DB	889	CGGATGAGCCAGTGGCCGAGGAGCCCTTACCTTCGCATGGAGTGGATGACCTACCTTA	948
QY	1081	AGGAGCGGCTGAAGAGGCTCATCTTCAGGAGACAGCAGCTTCAGGCCGAGAGTCTGG	1140
DB	949	AGGAGCGGCTGAAGAGGCTCATCTTCAGGAGACAGCAGCTTCAGGCCGAGAGTCTGG	1008
QY	1141	AGGCCCCCTAGCCACAGACAGCATCTCTGCACCCCTGGGGCTTGGAAACAGACTGGCAAG	1200
DB	1009	AGGCCCCCTAGCCACAGACAGCATCTCTGCACCCCTGGGGCTT-----	1050
QY	1201	AGGCCAAGAGGTCACTGAGGGCTCTGTCAACCAGGACCTGCTCCTGCTGCTGCCCTCTCC	1260
DB	1051	-----GGACCTGCTCCTGCTGCCCTCTCC	1077
QY	1261	CGCCAGACTGTTAGAAAATGGAACATGTGTCCAGCCCGGACCTTGGCGAGCCAGGCCGG	1320
DB	1078	CGCCAGACTGTTAGAAAATGGAACATGTGTCCAGCCCGGACCTTGGCGAGCCAGGCCGG	1137
QY	1321	GTGGAGCATGGGCTGGCCACCTCTCTCTTTGCTGAGGCTTCAGCTTCAGGCAAGGCA	1380
DB	1138	GTGGAGCATGGGCTGGCCACCTCTCTCTTTGCTGAGGCTTCAGCTTCAGGCAAGGCA	1197
QY	1381	AGGCTTCTCTCTCCACCCGCTTCCCAACGGGCTCTGGGAGCTCAGGTGGCCCCAGT	1440
DB	1198	AGGCTTCTCTCTCCACCCGCTTCCCAACGGGCTCTGGGAGCTCAGGTGGCCCCAGT	1257
QY	1441	TCATCTCCGCTGTCTGTCTGGGCTTACCTTCCAGCGCTCCAGTCTCTGGCAGT	1500
DB	1258	TCATCTCCGCTGTCTGTCTGGGCTTACCTTCCAGCGCTCCAGTCTCTGGCAGT	1317
QY	1501	TCGTGAATGGAAGGTTCTGGCTGCGCCCAACCTGCTGAAGGCGAGAGTGGAGGTGGGG	1560
DB	1318	TCGTGAATGGAAGGTTCTGGCTGCGCCCAACCTGCTGAAGGCGAGAGTGGAGGTGGGG	1377
QY	1561	GCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCAATCAAAACCCCAACCT	1620
DB	1378	GCGCTGAGTAGGAGCTCAGGGCCATGCTGCCCCCTCATCTCAATCAAAACCCCAACCT	1437
QY	1621	AGTTTCCCTGAAGGAACATTCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGCCG	1680
DB	1438	AGTTTCCCTGAAGGAACATTCTTAGTCTCAAGGCTAGCATCCCTGAGGAGCCAGCCG	1497
QY	1681	GSCCGAATCCCTCTCTGTCAAAGCTGTCACTTCGCTGCCCTCGCTGCTCTGTGTGTG	1740
DB	1498	GSCCGAATCCCTCTCTGTCAAAGCTGTCACTTCGCTGCCCTCGCTGCTCTGTGTGTG	1557
QY	1741	GTGAGCAGAGTGGAGCTGGGGGCGGTGGAGAGCCCGCGCCCTGCCACCTCCCTGACC	1800
DB	1558	GTGAGCAGAGTGGAGCTGGGGGCGGTGGAGAGCCCGCGCCCTGCCACCTCCCTGACC	1617
QY	1801	CGTCTAATATATAATATAGAGATGTGTATGGCTG	1837
DB	1618	CGTCTAATATATAATATAGAGATGTGTATGGCTG	1654

RESULT 6
US-11-060-756-1050
; Sequence 1050, Application US/11060756
; Publication NO. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1050
; LENGTH: 1400

781	CTGGGCTCCCATCCAGGAGGACCTGAAATTGATCATCAACATGAAGGCCCGAAACTAC	840
852	CTACAGTCTCTGCCCTCCAAAGACAAGGTGGCTTGGGCCAAAGCTTTTCCCAAGTCAGAC	911
841	CTACAGTCTCTGCCCTCCAAAGACAAGGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGAC	900
912	TCCAAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAAACCCCAATAAACCGGATCACA	971
901	TCCAAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAAACCCCAATAAACCGGATCA--	958
972	GTGGAGGAGCGCTGGCTCAACCCCTCACTGGAGCAGTACTATGACCCGCGGATGAGCCA	1031
959	-----CA	960
1032	GTGGCCGAGGAGCCCTTACCTTCGCCATGGAGCTGGATGACCTACTCTAAGGAGCGGCTG	1091
961	GTGGCCGAGGAGCCCTTACCTTCGCCATGGAGCTGGATGACCTACTCTAAGGAGCGGCTG	1020
1092	AAGGAGTCTCATCTTCCAGGAGACAGCACGCTTCCAGCCCGAGTGTCTGGAGGCCCTCTAG	1151
1021	AAGGAGTCTCATCTTCCAGGAGACAGCACGCTTCCAGCCCGAGTGTCTGGAGGCCCTCTAG	1080
1152	CCGACAGACATCTCTGCACCTTGGGGCTTGGACAG	1189
1081	CCGACAGACATCTCTGCACCTTGGGGCTTGGACCTG	1118

RESULT 11

```

US-10-072-036-56
; Sequence 56, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJORN
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mERK1-F64L-S65T-GFP fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1872)
; US-10-072-036-56

```

RESULT 12

US-10-171-311-126
; Sequence 126, Application US/10171311
; Publication No. US20030087270A1

```
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-171-311-126

Query Match      35.2%; Score 646.6; DB 5; Length 1611;
Best Local Similarity 73.2%; Pred. No. 1.6e-170;
Matches 829; Conservative 0; Mismatches 304; Indels 0; Gaps 0;

QY 1 GAGGATGAGATGGCGGCGCGCGGCTCAGGGGCGCGGGGCGGGAGCCCCGTAGAA 60
   |||||
Db 266 GAGGCGGCGCGCGGAGCGCTCGAGCTCAGCGCGCGGCGGCGCGCGCGGCGAGCCA 325

QY 61 CCGAGGGGTCGCGCGCGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCAGCGTTCG 120
   |||||
Db 326 ACATGCGGCGGCGGCGGCGGCGGCGCGCGCGCGGAGATGGTCCCGGGCAGGTTCG 385

QY 121 ACGTGGCGCGCGCTACAGCGAGTTCGAGTACATCGCGGAGGCGCGGTACGGCATGGTCA 180
   |||||
Db 386 ACGTGGGCGCGCTACACCAACTCTCGTACATCGCGGAGGCGCTACGCGATGGTCT 445

QY 181 GCTCGGCTATGACACCGTGGCGGCAAGCTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 240
   |||||
Db 446 GCTCTGTTATGATAATGTCAACAAAGTTCGAGTAGCTATCAAGAAATCAGCCCTTTG 505

QY 241 AACATCAGACCTACTGCCAGCGCAGCGCTCCGGGAGATCCAGATCCTGCTGCGTCCGCC 300
   |||||
Db 506 AGCACCAGACCTACTGCCAGAGAACCTGAGGGAGATAAAATCTTACTGCGCTTCAGAC 565

QY 301 ATGAGAAATGTCATCGCATCCGAGACATCTCGGGGCGTCCACCTCGGAAGCCATGAGAG 360
   |||||
Db 566 ATGAGAACATCATGGAATCAATGACATATTTCGACGACCAACCATCGAGCAATGAAG 625

QY 361 ATGCTTACATTTGCGAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAAGCCAGC 420
   |||||
Db 626 ATGTATATATAGTACAGGACCTCATGGAAACAGATCTTTTACAGCTCTTTGAAGACACAAC 685

QY 421 AGCTGAGCAATGACCATATCTGTACTCTCTTACCAGATCCTCGGGGCGCTCAAGTACA 480
   |||||
Db 686 ACCTCAGCAATGACCATATCTGTATTTTCTTACCAGATCCTCAGAGGTTTAAATATA 745

QY 481 TCCACTCCGCAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCA 540
   |||||
Db 746 TCCATTTCAGCTAACGTTCTGCACCGTGACCTCAAGCCTTCCAACCTGCTCTCAACACCA 805

QY 541 CTTGCGACCTTAAGATTGTTGATTTTCGGGCTGGCCCGGATTCGCCGATCTCTGAGCATGACC 600
   |||||
Db 806 CTTGTGATCTCAAGATCTGTGACTTTTGGCCCTGGCCCGGTGTTGCAGATCCAGACCATGATC 865
```

```
RESULT 13
US-10-301-822-114
; Sequence 114, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (328)...(1410)
; US-10-301-822-114
```

Query Match		35.2%; Score 646.6; DB 6; Length 1611;
Best Local Similarity		73.2%; Pred. No. 1.6e-170;
Matches		829; Conservative 0; Mismatches 304; Indels 0; Gaps 0;
Qy	1	GAGGAGTGGAGATCGCGCGCGCGGCTCAGCGGGGCGGGGCGGGAGCCCGTAGAA 60
Db	266	GAGCGCGCGCGCGCGGCTGCGAGCTCAGCGCGGAGGCGGCGCGCGCGCGCGCGCA 325
Qy	61	CCGAGGGGGTCCGGCGCGGGTCCCGGGGAGGTGGAGATGGTGAAGGGCAGCCGTTGC 120
Db	326	ACATGGCGCGCGCGCGCGCGCGCGCGCGCGGAGATGGTCCCGGGCAGGTGTTGC 385
Qy	121	ACGTGGGCGCGGTACACCGAGTTGCACTATCGCGGAGGGCGGTACGGCATGGTCA 180
Db	386	ACGTGGGCGCGGTACACCAACCTCTCGTACATCGCGGAGGGCGCTACGGCATGGTGT 445
Qy	181	GCTCGGCTATGACCGTGGCGGAAGTCTCGGTGGCCATCAAGAAGATCAGGCCCTTCG 240
Db	446	GCTCTGCTTATGATATGTCAACAAAGTTGAGTAGTATCAAGAAATCAGCCCTTTG 505
Qy	241	AACATCAGACCTACTGCCAGCGACGCTCCGGGAGATCCAGATCCTGCTCGCTTCGGCC 300
Db	506	AGCACAGACCTACTGCCAGAACCTTGAGGGAGATAAAATCTTACTCGCTTCAGAC 565
Qy	301	ATGAGAAATGTCATCGGCATCCGAGACATTCCTGGGGGGTCCACCCTGGAAGCCATGAG 360
Db	566	ATGAGAACATCATTTGAATCAATGACATTAATTCAGCACCAACCATCGAGCAAAATGAA 625
Qy	361	ATGCTACATTTGTCAGGACCTGATGAGACTGACCTGATCAAGTTGCTCAAAAGCCAGC 420
Db	626	ATGTATATATAGTACAGGACCTCATGGAACAGATCTTTTAAAGCTCTTGAAGACAAAC 685
Qy	421	AGCTGAGCAATGACATATCTGCTACTCTCTACAGATCTCTGGGGGCTCAAGTACA 480
Db	686	ACCTGACAAATGACCATATCTGCTATTTTCTACAGATCTCTCAGAGGGTTAAATATA 745
Qy	481	TCACCTCGGCAACGTCTCCACGAGATTAAGCCCTTCAACCTGCTCATCAACACCA 540
Db	746	TCCATTGAGTAACTTCTGACCGTCAAGCTTCAAGCTTCAACCTGCTCAACACCA 805
Qy	541	CTGCGACCTTAAGATTTGTGATTTCCGCTGGCGCGGATGGCGATCTGAGCATGACC 600
Db	806	CCTGTGATCTCAAGATCTGTGACTTTGGCTTGGCCCGGTGTGAGATCCAGACCATGATC 865
Qy	601	ACACCGGCTTCTGACGGAGTATGTGCTACGCGCTGATACCGGGCCCGAGAGATCATGC 660
Db	866	ACACAGGGTCTGACAGATATGTGCGCACAGTGTGTACAGGGCTCCAGAAATATGT 925
Qy	661	TGAATCCAAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCAATCTGGCTG 720
Db	926	TGAATCCAAAGGGCTACACCAAGTCCATTTGATATTTGGTCTGTAGGCTGCATTCGGCAG 985
Qy	721	AGATGCTCTTAACCGGCGCATCTTCCCTGGCAGCACTACTGGATCAGCTCAACACA 780
Db	986	AAATGCTTTCAACAGGCGCCATCTTTCCAGGGAAGCAATTTTGAACAGCTGAATCACA 1045
Qy	781	TTCTGGGCATCTCGGCTCCCATCCAGAGGACCTGAAATGTATCATCAACATGAAG 840
Db	1046	TTTTTGGGTATTTCTTGGATTTCCCATCAAGAAAGCTGAATGTATATATATTTAAAG 1105
Qy	841	CCCGAAACTACTACAGTCTCTGCCCTCCAGACCAAGGTGGCTTGGGCCAAGCTTTTCC 900
Db	1106	CTAGAACTATTTGCTTCTCTCCACACAAAATAAGGTGCCATGGAACAGGCTGTTC 1165
Qy	901	CCAGTACAGCTCCAAAGCCCTTGACCTGTGGACCGGATGTTAACTTTAAACCCCAATA 960
Db	1166	CAATGTGTGACTCCAAAGCTCTGGAATTTATTTGGAACAAAATTTGACATTTCAACCCACA 1225
Qy	961	AACGGATCAGATGGAGAGCGGTGGCTCAACCTTACCTGAGCAGTACTATCAACCCGA 1020
Db	1226	AGAGGATGAAGTAGAACAGGCTCTGCGCCCAACCATATCTGGAGCAGTATTAACGCCGA 1285
Qy	1021	CGGATGAGCCAGTGGCGGAGGCGCTTCACTTTCGCCATGGAGCTGGATGACCTACCTA 1080

Db	1286	GTGACAGGCCATCGCCGAGACCACTTCAAGTTCAACATGGAATGGATGACTTGCCTA 1345
Qy	1081	AGGAGCGGTGAAGGAGCTCATCTTCCAGGAGACACACGCTTCCAGCCCGGA 1133
Db	1346	AGGAAAAGCTAAAAGAACTAAATTTTGAAGAGACTGCTAGATTCAGCCAGGA 1398
RESULT 14		
US-10-305-720-1263		
; Sequence 1263; Application US/10305720		
; Publication No. US20040010136A1		
; GENERAL INFORMATION:		
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.		
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression		
; FILE REFERENCE: PA-0002-1 CON		
; CURRENT APPLICATION NUMBER: US/10/305,720		
; CURRENT FILING DATE: 2002-11-26		
; PRIOR APPLICATION NUMBER: 09/016,434		
; PRIOR FILING DATE: 1998-01-30		
; NUMBER OF SEQ ID NOS: 1490		
; SOFTWARE: PERL Program		
; SEQ ID NO 1263		
; LENGTH: 2791		
; TYPE: DNA		
; ORGANISM: Homo sapiens		
; FEATURE:		
; NAME/KEY: misc feature		
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g23878		
US-10-305-720-1263		
Query Match 34.9%; Score 642; DB 6; Length 2791;		
Best Local Similarity 74.9%; Pred. No. 3.6e-169;		
Matches 804; Conservative 0; Mismatches 270; Indels 0; Gaps 0;		
Qy	60	ACCGAGGGGTGCGCGCGGGTCCCGGGGAGGTGGAGATGTTGAAGGGGCGAGCCGTTTC 119
Db	96	ACCTGTGATGGCGCGCGCGCGGGCGCGGGCGCGGAGATGTTCCGGGGGAGGTGTTTC 155
Qy	120	GAGTGGGCGCGCTACACGCAAGTTGCAATATCATCGCGAGGGCGGCTACGGCATGGTC 179
Db	156	GAGTGGGCGCGCTACACCAACCTCTCGTACATCGCGAGGGCGCTACGGCATGGTG 215
Qy	180	AGTGGGCGCTATGACACAGTGGCGCAAGATCGCGGTGGCCATCAAGAGATCAGCCCTTC 239
Db	216	TGCTGCTTATGATAATGTCAACAAAGTTGCAAGTGTATCAAGAAATCAGCCCTTTT 275
Qy	240	GAACATCAGACCTACTGCGGAGCGCGCTCCGGAGATCCAGATCCTGCTGGCTTCGC 299
Db	276	GAGCACGAGACCTACTGCGGAGAGACCCCTGAGGAGATATAAATCTTACTGCGCTTCA 335
Qy	300	CATGAGAAATGTCATCGGCAATCCGAGACATTTCTCGGGCGTCCACCTTGGAGGCAATGAGA 359
Db	336	CATGAGAACATCATTTGGAATGACATTTATTCGAGACCAACCATCGAGCAAAATGAAA 395
Qy	360	GATGTCTACATTTGCGAGGACCTGATGAGAGCTGACCTGTACAGTTGCTGAAAGCCAG 419
Db	396	GATGTATATATAGTACAGGACCTCATGMAAACAGATCTTTTACAAGCTCTTTGAAGACAA 455
Qy	420	CAGTGGCAATGACCATATCTGCTACTTCTCTACAGATCCCTGCGGGCGCTCAAGTAC 479
Db	456	CACCTCAGCAATGACCATATCTGCTATTTCTCTACAGATCCCTCAGAGGTTTAAATAT 515
Qy	480	ATCCACTCCGCCAAGTGTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACC 539
Db	516	ATCATTCAGCTAACGTTCTGCAACGCTGACCTCAAGCTTCCAACTGCTGCTCAACACC 575
Qy	540	ACCTGCGACCTTAAGATTTTGTGATTTTCGGCTGGCCCGGATTCGCGATCTCTGAGCATGAC 599
Db	576	ACCTGTGATCTCAAGATCTGTGACTTTGGCGTGGCGGTGTCAGATCCAGACCATGAT 635
Qy	600	CACACCGGCTTCTGAGCGGAGTATGGCTACGGCTGTTACCGGGCCCGAGATCATG 659


```

Db      636  CACACAGGGTTCCTGACAGAAATATGTGGCCACACAGTTGGTACAGGGCTCCAGAAATATG 695
QY      660  CTGAACCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTCTGGCT 719
Db      696  TTGAATTCGAAGGGCTACACCAAGTCCATGATTAATTTGGTCTGTGGCTGCATTCTGGCA 755
QY      720  GAGATGCTCTCTAAACCGGCCATCTTTCCCTGGCAAGCCTACCTGATCAGCTCAACAC 779
Db      756  GAATGCTTTCTAACAGGCCATCTTTCCAGGAGACATTAATCTTACAGCTGAACAC 815
QY      780  ATTCTGGGATCTCTGGGCTCCCATCCAGGAGACCTGAATGTATCATCAACATGAAG 839
Db      816  ATTTTGGGTATCTTGGATCCCATCACAAAGACCTGAATGTATTAATAAATTTAAAA 875
QY      840  GCCGGAACCTACCTACAGTCTCTGCCCTCCAGACCAAGGTGCTTGGGCCAAGCTTTTC 899
Db      876  GCTAGGAACCTATTGTCTTCTTCCACACAAAATAAGGTGCTCAATGGAACAGGCTGTT 935
QY      900  CCCAAGTCCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACCCCAAT 959
Db      936  CCAATGCTGACTCCAAAGCTCTGGACTTATTTGGACAAAATGTTGACATTCACCCAC 995
QY      960  AAACGGATCACAGTGGAGGAGGGCTGGCTACCCCTACCTGAGGAGTACTATGACCCG 1019
Db      996  AAGAGGATTTGAAGTAGAACACAGGCTCTGGGCCACCATATCTGGAGCAGTATTACG 1055
QY      1020  ACGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTAC 1079
Db      1056  AGTGACGAGCCCATCGCCGAGACCATTTCAAGTTGAGATGGAAATGGATGACTTGCCT 1115
QY      1080  AAGGAGCGGCTGAAGGAGCTCATCTTTCCAGGAGACAGCACGCTTCCAGCCCGGA 1133
Db      1116  AAGGAAAGCTCAAGAACTAATTTTGAAGAGACTGCTAGATTCCAGCCAGGA 1169

RESULT 15
US-10-756-149-513
; Sequence 513, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 513
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-513

Query Match      34.9%; Score 642; DB 9; Length 2791;
Best Local Similarity 74.9%; Pred. No. 3.6e-169;
Matches 804; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

QY      60  ACCGAGGGGTCCGGCCCGGGGTCCCGGGGGAGGTGGAGATGGTGAAGGGCGCAGCCGTTTC 119
Db      96  ACCTCGTATGGCCGCGCGGGCGGGCGGAGATGGTCCGCGGCGAGGTGTTTC 155
QY      120  GAGTGGGCGCGCTACAGCGAGTTGACAGTACATCGGAGGCGGCGGTACGGCATGGTC 179
Db      156  GAGTGGGCGCGGCTACACCAACCTCTCGTACATCGGAGGCGGCGCTACGGCATGGTG 215
QY      180  AGCTCGGCTATCACACGTCGCCAGACTCGGTGCGCATCAAGAGATCAGCCCTTC 239
Db      216  TGTCTGCTTATGATATGTCACAAAGTTGAGTAGTACATCAAGAAATCAGCCCTTT 275
QY      240  GAACATCAGACCTACTGCCAGCGCAGCTCCGGGAGATCCAGATCCTGCTGGCTTCGCG 299

```

Search completed: February 7, 2006, 15:20:13
Job time : 1587.19 secs

```

Db      276  GAGCACACAGACCTACTGCGCAGAGAAACCCCTGAGGAGATAAAAAATCTTACTGGCTTCAGA 335
QY      300  CATGAGATGTTCATCGGCATCCGAGACATTCCTGCGGGGCTCCACCCCTGGAAGCCATGAGA 359
Db      336  CATGAGAACATCATTTGGGAATCAATGACATTAATTCGAGACCAACCATCGAGCAAAATGAAA 395
QY      360  GATGTCTCATATTGTGCAGGACCTGATGGAGACTGACCTGTGTACAGTTGCTGAAAAAGCCAG 419
Db      396  GATGTATATATAGTACAGGACCTCATGGAACAGATCTTTACAAGCTCTTGAAGACACAA 455
QY      420  CAGCTGAGCAATGACATATCTGCTACTCTACAGATCTCTGGGGGCTCTAAGTAC 479
Db      456  CACCTCAGCAATGACATATCTGCTATTTCTCTACAGATCTCTCAGAGGGTTAAAAATAT 515
QY      480  ATCCACTCCGCAACGCTGCTCCACAGATCTTAAAGCCCTCCAAAGCTGCTCATCAACACC 539
Db      516  ATCCATTCAGCTTAAAGCTTCTGCACCGTGACCTCAAGCCCTTCCAACTGCTGTCAACACC 575
QY      540  ACTCGGACCTTAAAGATTTGTGATTTCCGCGCTGGGCCCGGATTCGCCATCTCTGAGCATGAC 599
Db      576  ACTGTGATCTCAAGATCTGTGACTTTGGCCCTGGCCCGTGTTCAGATCCAGACCATGAT 635
QY      600  CACACCGGCTTCTGACGAGATGATGTGCTACGCGTGTGTACCGGGCCCGCAGAGATCATG 659
Db      636  CACACAGGGTTCTCTGACAGAAATATGTGGCCACACGTTGGTACAGGGCTCCAGAAATATG 695
QY      660  CTGAACCTCAAGGGCTTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTCTGGCT 719
Db      696  TTGAATTCGAAGGGCTTACCAAGTCCATTTGATTTGGTCTGTAGGCTGCATTCTGGCA 755
QY      720  GAGATGCTCTCTAAACCGGCCATCTTCCCTGGCAAGACACTACCTGGATCAGCTCAACCCAC 779
Db      756  GAAATGCTTTCTAAACAGGCCCATCTTCCAGGGAAGCATTAATCTTGACAGCTGAACCCAC 815
QY      780  ATTCTGGGATCCTGGGCTCCCATCCAGGAGGACCTGGAATTTGTATCATCAACATGAAG 839
Db      816  ATTTTGGGTATTTCTTGGATCCCATCACAAAGAACCTGAAATTTGATATAAATTTAAAA 875
QY      840  GCCGGAACCTACCTACAGTCTCTGCCCTCCAGACCAAGGTGCTTGGGCCAAGCTTTTC 899
Db      876  GCTAGGAACCTATTGTCTTCTTCCACAAAATAAGGTGCTCAATGGAACAGGCTGTTTC 935
QY      900  CCCAAGTCCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTAACCTTTAACCCCAAT 959
Db      936  CCAATGCTGACTCCAAAGCTCTGGACTTATTTGGACAAAATGTTGACATTCACCCACAC 995
QY      960  AAACGGATCACAGTGGAGGAGGGCTGGCTACCCCTACCTGAGGAGTACTATGACCCG 1019
Db      996  AAGAGGATTTGAAGTAGAACACAGGCTCTGGGCCACCATATCTGGAGCAGTATTACG 1055
QY      1020  ACGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCATGGAGCTGGATGACCTACCT 1079
Db      1056  AGTGACGAGCCCATCGCCGAGACCATTTCAAGTTGAGATGGAAATGGATGACTTGCCT 1115
QY      1080  AAGGAGCGGCTGAAGGAGCTCATCTTTCCAGGAGACAGCACGCTTCCAGCCCGGA 1133
Db      1116  AAGGAAAGCTCAAGAACTAATTTTGAAGAGACTGCTAGATTCCAGCCAGGA 1169

```


Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1708.4	93.0	1866	7	US-10-955-054A-5		Sequence 5, Appli
2	1059.8	57.7	1833	8	US-11-136-527-284		Sequence 284, App
3	646.6	35.2	1611	8	US-11-186-284-114		Sequence 114, App
4	630.8	34.3	1725	8	US-11-136-527-1988		Sequence 1988, App
5	275	15.0	1577	8	US-11-136-527-2485		Sequence 2485, Ap
6	273.4	14.9	3132	8	US-11-136-527-2850		Sequence 2850, Ap
7	273.2	14.9	1888	7	US-10-770-726-25		Sequence 25, Appl
8	260.8	14.2	2737	7	US-10-909-125-861		Sequence 861, App
9	233	12.7	1665	8	US-11-136-527-3828		Sequence 3828, Ap
10	225	12.2	1400	8	US-11-136-527-7924		Sequence 7924, Ap
11	203.8	11.1	3824	8	US-11-136-527-2977		Sequence 2977, Ap
12	166.4	9.1	822	8	US-11-136-527-3508		Sequence 3508, Ap
13	185.8	9.0	1980	8	US-11-136-527-238		Sequence 238, App
14	165	9.0	600	8	US-11-136-527-7604		Sequence 7604, App
15	164.2	8.9	600	8	US-11-136-527-4380		Sequence 4380, Ap
16	147.8	8.0	900	8	US-11-136-527-55		Sequence 55, Appl
17	141.8	7.7	3131	7	US-10-770-726-28		Sequence 28, Appl
18	136.2	7.4	2765	8	US-11-136-527-126		Sequence 126, App
19	135.8	7.4	2804	8	US-11-136-527-282		Sequence 282, App
20	135.4	7.4	2211	8	US-11-127-817-10		Sequence 10, Appl
21	135.4	7.4	2372	7	US-10-857-780-11		Sequence 11, Appl
22	135.4	7.4	2372	8	US-11-127-817-9		Sequence 9, Appli

182 CTGGGCTATGACCACTGCGGCAAGACTCGCGTGGCCATCAAGAAGATCAAGCCCTTCGA 241
198 CTGAGCWTATGACACAGTGGCGAAGACCAAGTGGCTATCAAGAAGATVAGCCCTTCGA 257
242 ACATCAGACCTACTGCGGCGGAGGCTCGGAGATCCAGATCCTGCTGCGCTTCGCCA 301
258 GCATCAAACTACTGTCAGGSCAGCTGAGAGAAATCCAGATTTGCTSSGATTCGCCA 317
302 TGAGATGTCTATCGGCAATCGAGACATTTCTGGGGGCTCCACCTCGAAGCCATGAGAGA 361
318 TGAGATGTCTATAGGCTATCGAGACATCTTCAGAGACCCACCTCGAAGCTATGAGAGA 377
362 TGTCTACATTTGTCAGGACCTGATGGAGACTGACCTGTACAAGTGTCTGAAAAGCCAGCA 421
378 TGTTTACATTTGTCAGGACCTGATGGAKACGACCTGTACAAGCTGTAAAGGCCARCA 437
422 GCTGAGCAATGACATATCTGCTACTTCTCTACAGATCCTCGGGGCTCAAGTACAT 481
438 GCTGAGCAATGACATATCTGCTACTTCTCTACAGATCCTCGGGGCTCAAGTACAT 497
482 CCACTCGGCAACCTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCAC 541
498 AACTCGGCCAATGTGCTGACCGGACCTGAGGCCCTCCAACTGCTCATCAACACCAC 557
542 CTGCGACCTTAAGATTTGATTTGGCTGCGCCCGGATTCGCGATCTCTGAGCATGACCA 601
558 CTGAGCCTTAAGATCTGATTTGGCTGCGCCCGGATTCGCGATCTCTGAGCATGACCA 617
602 CACGGCTTCTGACGAGATGATGGCTACGCGTGTGACCGGCCCCAGAGATCATGCT 661
618 CACTGGCTTTCTKACCGAGTATGTGGCCACACGCTGTACCGAGCCCCAGAGATCATGCT 677
662 GAATCCAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGACATCTGGCTGA 721
678 TAATCCAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGACATCTGGCTGA 737
722 GATGCTCTTAACCGGCCCCATCTTCCCTGGCAAGCACTACTCGATCAGCTCAACCAT 781
738 GATGCTCTTAACCGGCCCCATCTTCCCTGGCAAGCACTACTCGATCAGCTCAACCAT 797
782 TCTGGGATCTGCGGCTCCCGATCCAGAGGAGCTGAATGATCATCAACATGAGGC 841
798 TCTAGTATATCTGGTTTCCCATCCCAAGAGGACTTAAATGATCATTAACATGAGGC 857
842 CCGAAGCTACTCAGTCTGCGCTCCCAAGCAAGTGGCTGGGCCCAAGCTTTTCCC 901
858 CCGAAGCTACTCAGTCTGCGCTCCCAAGCAAGTGGCTGGGCCCAAGCTTTTCCC 917
902 CAAGTCAAGCTCCAAAGCTTGAACCTGGAGCGGATGTTAACTTTAAACCCCAATAA 961
918 CAATCTGACTCCAAAGCTTGAACCTGGAGCGGATGTTAACTTTAAACCCCAATAA 977
962 ACGATCAGAGTGAAGAGCTGATCTTCCAGGAGACAGCACTTCCAGCCCGGAGTGTGGA 1021
978 GCGCATCAGAGTGAAGAGCTGATCTTCCAGGAGACAGCACTTCCAGCCCGGAGTGTGGA 1037
1022 GGATGAGCAGTGGCCGAGGAGCCCTTCACTTCCGATGAGCTGATGATGATGATGATGAT 1081
1038 AGATGAACAGTGGCTGAGGAGCCATTCACCTTTGATGATGAGCTGATGATGATGATGAT 1097
1082 GGAGCGGCTGAAGAGCTGATCTTCCAGGAGACAGCACTTCCAGCCCGGAGTGTGGA 1141
1098 GGAGCGGCTGAAGAGCTGATCTTCCAGGAGACAGCACTTCCAGCCCGGAGTGTGGA 1157
1142 GGCCCTTACGAGCAGACATCTGTCACCTTGGGGCTGGAAAGAACTGGCAAGA 1201
1158 GGCCCTTACGAGCAGACATCTGTCACCTTGGGGCTGGAAAGAACTGGCAAGA 1200
1202 GGCAAGAGTCACTGAGGGCTCTGTACCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1261
1201 -----TGTGCTTACCTGCTCTCTC

1262 GCAGACTTTAGAAAATGGAACACTGTGCCCCGAGCCGACCTTGGCAGCCCGAGCCGGG 1321
1223 TGCAGATTTGTAGAAAATG--AACTTTGTCTAAACCGGACCCCGGAGCCCGAG--CTGA 1279
1322 TGAGCATGGGCTGCGCACCTCTCTCTTTTGTGAGGGCTCTCAGCTTTCAGGCGAGCCAA 1381
1280 CCNAGGGTGGGCTGGCACCCCTCTCCTGCTGCTGGGGTCTCTCTTCAAG--AGGCTT 1337
1382 GGCCTTCTCTCCCAACCCGCTTCCCAAGGGGCTTCGGGAGCTCAGTGGCCCAAGTT 1441
1338 CTCCCACTCCAGTCCCTGCCCCATCTCCCTTGAAGCTGATGAGTGGTCCAGAG 1397
1442 CAATCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1501
1398 CTGATCTGCTGCTG--TGTCTTTTATCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 1456
1502 CTGGAATGAAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1561
1457 CTGGAATGAAGGGCTATGACCCCTA-----GGACCTGTCTACAGAGGGTGGAGG 1510
1562 GGCCTGAGTAGGACTCAGGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1621
1511 GCACTGAGT-----AGGCTAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1562
1622 GTTTCCTGAGGAACATTCCTTAGTCTCAAGGGCTAGCATCCCTGAGGAGCCAGGCCGG 1681
1563 TTTTCCTGACAGAACTTCTTAAATCTCAAGGGCTAGTCTTCTGAGGAGCCAGCCTAG 1622
1682 GCGAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1623 GCTAACCTCTCTCTCTC--AGCTGCCACATGTAAGCCCTGCTGCTGCTGCTGCTGCTGCT 1681
1741 GTGAGCAGAGTGGAGCTGGGGGCTGGAGAGCCCGGGGCTGCTGCTGCTGCTGCTGCTGCT 1800
1682 GTGATGGATGTGGAGGCGGGGCTGGAGAGCCCGGGGCTGCTGCTGCTGCTGCTGCTGCT 1741
1801 C---GCTTAATATAATATAGAGATGCTCTATGG 1834
1742 CTGATCTAATATAATATAGAGATGCTATATGG 1778

RESULT 3

US-11-186-284-114
; Sequence 114, Application US/11186284
; Publication No. US2005026493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burt, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 1611
; TYPE: DNA

```

; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (328)...(1410)
; US-11-186-284-114

Query Match      35.2%; Score 646.6; DB 8; Length 1611;
Best Local Similarity 73.2%; Pred. No. 4.6e-139; Indels 0; Gaps 0;
Matches 829; Conservative 0; Mismatches 304;

QY 1 GAGGAGTGGAGATGGCGGGCGGGCTCAGGGGGCGGGGGCGGGGAGCCCCGTAGAA 60
Db  GAGGCGCGCGCGCGAGCGTCGAGCTCAGCGGGCGGAGCGGCGCGCGCGCGCGCGCA 325
QY 61 CCGAGGGGTGCGCCCGGGGGTCCCGGGGAGGTGGAGATGGTGAAGGGGCGAGCGGTTCG 120
Db  ACATGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGAGATGGTCGCGGGCGAGGTGTCG 385
QY 121 ACGTGGGCGCGGCTACACGCGAGTTCAGTACATCGCGGAGGCGGCTACGGCATGGTCA 180
Db  ACGTGGGCGCGGCTACACCAACTCTCTGTACATCGCGGAGGCGGCTACGGCATGGTGT 445
QY 181 GCTCGGCCCTATGACCGCGTGGCAAGCTCGCGTGGCCATCAAGAAGATCAGCCCTTCG 240
Db  GCTCTGCTTATGATAATGTCAACAAAGTTCGAGTAGCTATCAAGAAATCAGCCCTTTG 505
QY 241 AACATCAGACCTACTCCGAGCGACGCTCGGGAGATCCAGATCTCTGCTGGCTTCGGCC 300
Db  AGCACCAGACCTACTCCGAGAGAACCTCGGGAGATCAAAATCTTACTGCGCTTCAGAC 565
QY 301 ATGAGAAATGTATCGGCATCCGAGACATCTCGGGCGCTCCACCTTGGAGCATGAGAG 360
Db  ATGAGAAATCATTTGGAATCAATGACATTTTCGAGCACCAACCATCGAGCAATGAAG 625
QY 361 ATGCTTACATTTGTCAGGACCTGATGGAGATGACCTGTACAAAGTTGCTGAAAGCCAGC 420
Db  ATGTATATATAGTACAGGACCTCATGAAACAGATCTTTTACAAGCTCTTTGAAGACACAAC 685
QY 421 AGCTGAGCAATGACCAATATCTGCTACTTCTCTACAGATCTCTGCGGGGCTCAAGTACA 480
Db  ACCTCAGCAATGACCAATATCTGCTATTTCTCTACAGATCTCTCAGAGGTTTAAATATA 745
QY 481 TCCACCTCCGCAACGTCCTCCACGAGATCTAAAGCCCTCCACCTGCTCATCAACACCA 540
Db  TCCATTCAGTACAGTCTGCAACCGTGACCTCAAGCCTTCCAACTGCTGCTCAACACCA 805
QY 541 CTTGCGACCTTAAAGTTTGTGATTTTCGGCTTGCGCGGATGTCGATCTGAGCATGACC 600
Db  CCTGTGATCTCAAGATCTGTGACTTTTGGCTTGCGCGCTGTTGCGATCCAGACCATGATC 865
QY 601 ACACCGGCTTCTGACGGAGTATGTCGCTACGGCTGAGTACGGGGCCCGAGAGATCATGC 660
Db  ACACAGGGTTCCTGACAGAAATATGTGGCCACACGTTTGGTACAGGGCTCCAGAAATATGT 925
QY 661 TGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGCTGTGGGCTGCATTTCTGGCTG 720
Db  TGAATTTCAAGGGCTACACCAAGTCCATGTATATTTGGTCTGTAGGCTGCATTTCTGGCAG 985
QY 721 AGATGCTCTCTAAACCGGCCCATCTTCCCTGGCAAGCACTACCTTGGATCAGCTCAACACA 780
Db  AATATGCTTTCMAACAGGCCCATCTTTCAGGGAAGCAATATCTTTCAGCCAGCTGAATCACA 1045
QY 781 TTCTGGGCATCTGGGCTCCCATCCGAGGAGACCTGAATGTATCATCAACATCAAGS 840
Db  TTTTGGGTATTTCTTGGATCCCCATCAAGAAGACCTGAATGTATTAATAATTTTAAAG 1105
QY 841 CCGGAAATACCTACAGTCTCTGCCCTCCAAAGACCAAGGTGGCTTGGGCCCAAGCTTTTCC 900
Db  CTAGGAACATAATTTGCTTTCTCTTCCACACAAATTAAGGTGCCATCGAACAGGCTGTTC 1165
QY 901 CCAAGTCAGATCCAAAGCCCTTGACCTGCTGGAACGGATGTTTAACTTTTAAACCCCAATA 960
Db  CAAATGCTGACTCCAAAGCTCTGGAATTTATTTGACAAATAATGTTGACATTCACCCACACA 1225

```

```

QY 961 AACGGATCACAGTGGAGGAAGCGCTGGCTCACCCCTACTCTGGAGCAGTACTATGACCCGA 1020
Db  AGAGGATTTGAAGTAGACAGAGCTCTGGCCCAACCCATATCTGGAGCAGTATTTAGACCCGA 1285
QY 1021 CGGATGAGCCAGTGGCCGAGGAGCCCTTCACTTCGCCCATGGAGCTGGATGACCTTACCTA 1080
Db  GTGACGAGCCCATCGCGAAGCACCATTTCAAGTTCGACATGGAAATGGATGACTTGCCTA 1345
QY 1081 AGGAGCGGCTGAGAGGCTCATCTTCCAGGAGACAGCAGCGTTCCAGCCCGGA 1133
Db  AGGAAAGACTAAAGAACTAAATTTTGAAGAGACTGTAGATTTCAGCCGAGGA 1398

RESULT 4
US-11-136-527-1988
; Sequence 1988, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AMI01086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1988
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-1988

```

```

Query Match      34.3%; Score 630.8; DB 8; Length 1725;
Best Local Similarity 74.5%; Pred. No. 2e-135;
Matches 788; Conservative 4; Mismatches 266; Indels 0; Gaps 0;

QY 76 CCGGGGTCCCGGGGAGGTGGAGATGTTGAAGGGGCGAGCCGTTCCAGCTGGGCGCCGCGCT 135
Db  CCGGGGCGGCGCGCGCGCGCGGAGATGGTCCGCGGCGAGGTGTTCCAGCTGGGCGCGCGCT 238
QY 136 ACACGCGAGTTGCGATACATCGCGGAGGCGGCTACGGATGCTCAGCTCGGCTATGACC 195
Db  ACATTAATCTCTGATACATCGGAGAGGCGCTACGGCATGTTGTTCTGCTATGATA 298
QY 196 ACGTGGCAAGACTCGGTGGCCATCAAGAAAGATCAGCCCTTCGAAACATCAGACCTACT 255
Db  ATCTCAACAAAGTTCGAGTTGCTATCAAGAAATCAGTCTTTTGAGCACACGACCTACT 358
QY 256 GCGAGGCGACGCTCGGGGAGATCCAGATCCTCGCTTCCGCCATGAGAAATGTCATCG 315
Db  GTCCAGAGAACCTGAGAGAGATAAAATCTTACTGCGCTTTCAGACATGAGAAATCATCG 418
QY 316 GCATCCGAGACATCTGCGGGCGCTCCACCTCGGAAGCCATGAGAGATGTTCTACATTGTC 375
Db  GCATCAATGACATCATCCGGGACCAACCATTTGAGCAGATGAAGATGATATATAGTAC 478
QY 376 AGGACCTGATGAGACTGACCTGTACAAGTTCTGTAAGAGCCAGCAGCTGAGCAATGACC 435
Db  AGGACCTCATGAGAGACAGATCTTTTCAAGCTCTTGAAGACACAGCAACCTTCAGCAATGATC 538
QY 436 ATATCTGCTACTTCTCTTACAGAGATCCTGCGGGGCTCAAGTACATCCACTCCGCCAAG 495
Db  ATATCTGCTATTTCTTTTATCAGATCCTGAGAGGATTTAAAGTATATACATTCAGCTAATG 598
QY 496 TGTCTCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACACCACTCGGACCTTTAAGA 555
Db  TTTCTGACCGTGAACCTCAAGCCTTCCAACTCTCTGCTGAACCACTTGTGATCTCAAGA 658
QY 556 TTTGTGATTTTGGCTTGGCCCGGATTTGCGGATCTCTGAGATCAACCAACCGGCTTCTGA 615

```

Db 659 TCTGTGACTTTGGCGCTTGGCCGTTGTCAGATCCAGATCCAGATCATACAGGGTCTTGA 718
Qy 616 CGAGATGATGGCTTACCGCGTGTGTTACCGGGCCCGAGAGATCATGCTGAACTCCAAGGGCT 675
Db 719 CAGATATGTAGCACACGGTGGTACAGAGCTCCAGAAATATGTTGAATTTCCAAGGGTT 778
Qy 676 ATACCAAGTCATCGACATCTGCTGTGGCTGTCATCTTGGCTGAGATGCTCTTAACC 735
Db 779 ATACCAAGTCATGATATTTGGTCTGTGGCTGTCATCTTGGCAGAGATGCTATCCAACA 838
Qy 736 GGCCCATCTTCCCGGCAAGCACTTACCTGGATCAGCTCAACCAATTTCTGGGATCCTGG 795
Db 839 GGCCTATCTTCCCGGAAAGCAATTTACCTTGACAGCTGAAATCAATCTCGGTAATCTTG 898
Qy 796 GCTCCCCATCCAGAGAGACTGAATTTGTATCATCAACATGAAGGCCCGAACTTACCTAC 855
Db 899 GATCTCCATCACAGAGAGATCTGAATTTGTATATAAATTTAAAGCTTAGAAACTATTTC 958
Qy 856 AGTCTCTGCCCTCCAGACCAAGTGGCTTGGGCAAGCTTTTCCCAAGTCAAGTCCA 915
Db 959 TTTCTCTCCGCACAAAAATAAGTGGCGGTGGAACAGGTTGTTCCCAAGCGCTGACTCCA 1018
Qy 916 AAGCCCTTGACTCTGCGACCGGATGTTAACTTTTAAACCCCAATAACGGATCACATGG 975
Db 1019 AAGCTCTGGAATTTACTGGATAAAATGTTGACATTTAAACCTTCAAGAGGATGAAGTG 1078
Qy 976 AGGAAGCGCTGGCTCACCCCTTACCTGAGAGAGTACTATGACCCGAGATGAGCCAGTGG 1035
Db 1079 AACAGGCGKRGSCCACCCTGATCTCTGAGCAGTATTAATGACCCAAAGTATGAGCCATTG 1138
Qy 1036 CCGAGAGCCCTTCCACTTCCGCTTGGAGTGGATGACCTTCAAGAGCGGCTGAAG 1095
Db 1139 CTGAAGCACCATTCAAGTTTGACATGAGCTGGAGGACTTACCTTAAGGAGAGCTCAAAG 1198
Qy 1096 AGCTCATCTTCCAGAGACAGCAGCTTCCAGCCCGGA 1133
Db 1199 AACTCAATTTTGAAGAGACTGCTCGAATTCAGCCAGGA 1236

RESULT 5

US-11-136-527-2485
; Sequence 2485, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2485
; LENGTH: 1577
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2485

Query Match 15.0%; Score 275; DB 8; Length 1577;

Best Local Similarity 56.9%; Pred. No. 9.1e-54;

Matches 597; Conservative 0; Mismatches 425; Indels 27; Gaps 4;

Qy 115 CGTTGAGCTGGGCGCGCTTACAGCAGTGTGAGTATCATCGGAGGCGCGCTACGGCA 174
Db 61 CTGGGAGCTGCCAAGACTTACCTTGGCGCCCGGACGCTCGGAGCGGGCGCTATGGCG 120
Qy 175 TGGTCAGCTCGGCTATACCAAGCTGGCCAGACTCGGCTGCCATCAAGAGATCAGC- 233
Db 121 CGGTGTGCTCGGCCATCAGCAAGGGAAGGAGAGAGTGGCCATCAAGAGCTGAGCC 180
Qy 234 --CCCTCGAATACAGACTTACTGCCAGCGACGCTCCGGGAGATCCAGATCTCTGTGTC 291

Db 181 GGCCCTTCCAGTCCGAGATCTTTGCCAAGCGCGCTTACCGGAACTCTCTGCTGTGAAGC 240
Qy 292 GCTTCGCGCCATGAGATGTATCGGCATCCGAGA----CATTTGCGGGGGTCCACCCTGG 348
Db 241 ACATGACCATGAGAACGTCATTGGGCTTCTGATGTCTACACACCTGCCACTTCCGTTTC 300
Qy 349 AAGCCATGAGAGATGTCTATATTGTGAGAGACCTGATGAGAGCTGACCTGTACAAAGTTC 408
Db 301 GAAACTTCCAAGATTTTCTACCTGTTGATGCTTTTCAATGACAGACCGACCTTCGAGAATAA 360
Qy 409 TGAAGAGCCAGCAGCTGAGCAATGACCATATCTGCTACTTCTCTTACAGATCTCGGGG 468
Db 361 TG---GGGATGAGTTCAGTGAGAGAGAGGTCCAGTATTTGGTGTACCAAGTCTCAAAG 417
Qy 469 GCCTCAAGTACATCCACTCTCGCCAAAGTGTCTCACCGAGATCTAAAGCCCTTCAACCTGC 528
Db 418 GTCTAAGTACATCCACTCAGCTGGTATCGTTCAAGGGACCTGAAAGCCAGGCAACTGG 477
Qy 529 TCATCAACACCACTCTGGACCTTAAAGATTTGTGATTTGCGCTGGCCCGGATTCGCCATC 588
Db 478 CCCTGAATGAAGACTGTGAGCTGAAGATCTCGACTTTGGGCTGGCAAGCCACACGATG 537
Qy 589 CTGAGCATGACCACACCGGCTTCTGACGAGATATGTGGCTACGCGCTGTACCGGGCCC 648
Db 538 CGGAG-----ATGACTGGCTATGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGT 579
Qy 649 CAGAGATCATGCTGAATCTCAAGGGCTATACCAAGTCCATCGACATCTGGTCTGTGGGCT 708
Db 580 CTGAGGTGATCTCAGCTGGATGATTAACACAGAGCGGTGGATATCTGTGCTGTGGCT 639
Qy 709 GCATCTGGCTGAGATGCTCTTAAACGGCCCATCTTTCCCTGGCAAGCACTACTCTGATC 768
Db 640 GCATCATGGCCGAGATGCTTAAACGGAAAGACACTCTTCAAGGGGTAAAGGACTTACCTG 699
Qy 769 AGCTCAACACACTTCTGGGCTCTCTGGGCTCCCATCCAGAGGAGCTCGAATTTGTATCA 828
Db 700 AGCTGACCCAGATCTCTGAAGTGAAGTGAAGTGGGTTGCGAGTGGCGAGTTCGTGAGAGCTGA 759
Qy 829 TCACATGAAGGCGCGGAACTTACCTACAGTCTCTGCGCTTCCAGAGCAAGGTGGCTTGGG 888
Db 760 AAGCAAGGCGGCGCAATCTTACATTCAGTCTCTGCGCCAGAGCGCCCAAGAGGATTTCA 819
Qy 889 CCAAGCTTTTCCCAAGTGAAGTCTTCAAGCCCTTGAAGCTGCTGGACCGGATGTTACCT 948
Db 820 CACAGCTTTTCCCAAGGCGCGGCGCAAGCTGTAGACCTGCTGGAAGAATGCTGAGC 879
Qy 949 TTAACCCCAATAAAGGATCAGAGTGGAGGAGCGCTGCTCACCCCTTACCTGAGCAGT 1008
Db 880 TGGACGTGGACAAGCGCTTACCGCTCTCAGGCACTCGCTCACCCCTCTTTGAAACCCC 939
Qy 1009 ACTATGACCCGAGCGATGAGCCAGTGGCCGAGGAGCGCTTTCACCTTCGCCATGAGCTGG 1068
Db 940 TCCGGAGCCCTGAGGAGGAGACAGAGGCCCGCAGCAGCATTTGATGATGCTTAGAGCGGG 999
Qy 1069 ATGACCTTACCTAAGGAGCGGCTGAAGAGTGTATTTTCCAGGAGACAGAGCTTCCAGC 1128
Db 1000 AGAATCTCAGCGTGGAGCAATGGAACAACACATCTACAAAGAGATCGCCAACTTCAGTC 1059
Qy 1129 CCGGAGTGTGGAGGCGCCCTTAGCCCA 1157
Db 1060 CCATAGCCGGAAGGACTCGAGCGCAGCA 1088

RESULT 6

US-11-136-527-2850
; Sequence 2850, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)

; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2850
; LENGTH: 3132
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-11-136-527-2850

Query Match 14.9%; Score 273.4; DB 8; Length 3132;
Best Local Similarity 57.2%; Pred. No. 2.7e-53;
Matches 544; Conservative 11; Mismatches 372; Indels 24; Gaps 3;

QY 95 GGAGATGGTGAAGGGCAGCCGCTTCACGCTGGGCGCGCTACACGCAATGTCAGTACAT 154
DB 41 GCAGGAGCTGAACAAGACCGCTCTGGGAGGTGCCGAGCGATACCAGAACCTGTCCCGGT 100

QY 155 CGCGAGGGCGGTACGGCATGCTCAGCTCGGCTATGACCACTGCGGCAAGACTCGCT 214
DB 101 GGGCTCGGAGCTACGGCTCGGTGTGTGCTGTWYGAACCAAGACGGGACATCGTGT 160

QY 215 GGCATCAAGAAGAT---CAGCCCTTCGAAACATCAGACTACTGCCAGCGACGCTCG 271
DB 161 GGCAGTGAAGAGCTCTCGAGCCGCTTTCAGYCCATCATTCAGCCAAAGGWCCTACAG 220

QY 272 GGAGATCCAGATCCTGCTCGGCTTCGCGCATGAGAAATGTCATCGGCATCCGAGACATTC 331
DB 221 GGAGCTCGGCTGCTGAAGCACAATGAACACGAGAAATGTGATGGTCTGTGGATGTGT 280

QY 332 CGGGCGTCCA---CCCTGGAGCCATGAGAGATGCTACATTTGTCAGGACTGATGGA 388
DB 281 TACACCTGCAAGGTCCCTGGAGAAATTCACGATGTGTACCTGGTACCCTCTCATGG 340

QY 389 GACTGACCTGTACAACTTGTCTGAAAGCCAGCAGCTGAGCAATGACCATATCTGTACTT 448
DB 341 GGCAGACCTTGAAACAAATCGTGAAGTGTGAGAGCTTACCGATGACCAAGTTTCAGTTCT 400

QY 449 CTTTACCAAGATCCTCGGGGCTCAAGTACATCCATCGCAACGTCGCTCCACCGAGA 508
DB 401 TATCTACCAAGATCCTCGAGGGCTGAAGTATATACACTCGGCTGACATAATCCACAGGA 460

QY 509 TCTTAAGCCCTCAAACCTGCTCATCAACACCACCTCGGACCTTAAGATTTGATTTCCG 568
DB 461 CTTAAAGCCCAAGCAACCTCGCTGAATGAAGACTGTGAGCTGAAGATTTCTGGATTTGG 520

QY 569 CTGGCGCCGATTCGCGATCCTGAGCATGACCACACCGGCTTCCTGACGGAGTATGGC 628
DB 521 GCTGGCTCGGCACACTGAT-----GAGGAAATGACCGGCTACGTGGC 562

QY 629 TACGCGCTGGTACCGGCCCCAGAGATCATGTGAACTCAAAGGGCTATACCAAGTCCAT 688
DB 563 TACCCGGTGGTACAGAGCCCCGAGATATGCTGAATTTGATGCACTACAAACAGACAGT 622

QY 689 CGACATCTGGTCTGTGGCTGCAATCTGGCTGAGATGCTCTTAAACGGGCGCATCTTCC 748
DB 623 GGATATTTGGTCCGTGGCTGCAATCATGCTGAGCTGTGACCGGAAGAACGTTGTTTCC 682

QY 749 TGGCAAGCACTACTCGATCAGCTCAACACATTTCTGGGCATCTCGGGCTCCCCATCCCA 808
DB 683 TGGTACAGACCATATTGATCAGTTGAAGCTCATTTTAAGACTGTTGGAACCCCGGGGC 742

QY 809 GGAGGACCTGAATTTGATCATCAACATGAAGGCCCGCAAACTACCTACAGTCTTTCGCCCTC 868
DB 743 TGAGCTTCTGAAGAAAATCTCTCTCAGAGTMTGCAAGAAAATCATTTCACTCTTCGCCCA 802

QY 869 CAAGACCAAGTGGCTTGGGCCAAGCTTTTCCCAAGTCAGCTCAGAAAGCCCTTCACCT 928
DB 803 GATGCCGAAGATGAATTCGCAAAATGTATTTATTTGGTGCAATTCCTCTGGCTGTCACCT 862

QY 929 GCTGGAGCCGATGTTAACTCTTTAAACCCCAATAAACGATACAGTGGAGGAGCGTGGC 988

DB 863 GCTGCAAAAGATGCTGGTTTTTGGACTCRGATAAGAGGAWCACAGCAGCCCAAGCTCTTGC 922
QY 989 TCACCCCTACCTGGAGCAGTACTATGACCCGACGATGAGCCAGTGGCGGA 1039
DB 923 GCATGCCTACTTTTGTCTCAGYACCAACGACCCCTGATGATGAGCCAGTGGCTGA 973

RESULT 7

US-10-770-726-25
; Sequence 25, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Brown, Eugene

; APPLICANT: Liu, Wei

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: AM101079 (031896-010000)

; CURRENT APPLICATION NUMBER: US/10/770,726

; CURRENT FILING DATE: 2004-02-04

; NUMBER OF SEQ ID NOS: 48640

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 25

; LENGTH: 1888

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-770-726-25

Query Match 14.9%; Score 273.2; DB 7; Length 1888;

Best Local Similarity 56.5%; Pred. No. 2.5e-53;

Matches 603; Conservative 0; Mismatches 438; Indels 27; Gaps 4;

QY 95 GGAGATGTGAAGGGCAGCCGTTGACGCTGGGCGCGCTACACGCAATGTCAGTACAT 154
DB 131 GCAGGACGTCACAAAGACCGCTCGGAGCTGCCAAGACTTACGTGTCCCCGACGACGT 190

QY 155 CGCGAGGGCGGCTACGGCATGCTCAGCTCGGCTTATGACCACTGTCGCGCAAGACTCGCT 214
DB 191 CGGACAGGGGCTATGGCTCGCTGCTCGGCCATCGACACGCTCAGGGGAGAAAGT 250

QY 215 GGCCTCAAGAAGATCAGC---CCCTTGCAACATCAGACTTACTGCGCAGCGACGCTCG 271
DB 251 GGCCTCAAGAAGCTGAGCCGACCCCTTTCACTCCGAGATCTTCGCCAAGCGGCGCTACCG 310

QY 272 GGAGATCCAGATCCTGCTCGCTTCGCGCATCAGATGTCTACGCTCGCATCCGAGA---CAT 328
DB 311 GGAGCTGTCTGCTGTAAGCAATGAGCATGAGAACGTCATGAGAACGTCATTGGCTCTGATGTC 370

QY 329 TCTGGGGCGCTCCACCCTGGAAGCCATGAGAGATGTCTACATTGTGCGAGACCTGATGGA 388
DB 371 CACCCAGCGCTCTCCCTGGCAACTTCTATGACTTCTACCTGCTGATGCCCTTCATGCA 430

QY 389 GACTGACCTGTACAGTTGCTGAAAGCCAGCAGCTGAGCAATGACCATATCTGTACTT 448
DB 431 GACCGATCTGCAAGAAGATCATG---GGATGAGTTTCACTGAGAGAGATCCAGTACCT 487

QY 449 CTTTACCAAGATCCTGGGGGCTCAAGTACATCCACTCCGCCAAGCTGCTCCACCGAGA 508
DB 488 GGTGATCAGATGCTCAAGGGCTTAAAGTACATCACTCTGCTGGGTCGTGCAACGGGA 547

QY 509 TCTAAAGCCCTCCAACTCTCTCATCAACACCACTGCGACCTTAAAGATTGTGATTCGG 568
DB 548 CTGGAAGCCAGCAACCTGGCTGTGAATGAGGACTGTGAACCTGTAAGATTCTGATTTGG 607

QY 569 CTGGCGCGGATGCGGATCTCTGAGCATGACACCCGCTTCTCTGACGGAGTATGTGSC 628
DB 608 GCTGGCGGCATGTCAGACGCGCAG-----ATGACTGGCTACGTGTGT 649

QY 629 TACGCGCTGGTACCGGCCCCCAGAGATCATGCTGAACCTCAAGGGCTATACCAAGTCCAT 688
DB 650 GACCCGCTGGTACCAGGCCCCCGAGGTGATCTCAGCTGGATGCATCAACACAGACGT 709

Qy	739	CCATCTTCCCTGGCAAGCACTACTCTGGATCAGCTCAACCACTTCTGGGCATCTCTGGCT	7398
Db	586	TCCTGTTTCAAAGGCAATGACCACTGGACAGCTGAAGGAGATCATGAAAGTCACAGGA	645
Qy	799	CCCCATCCCCAGGAGGACCTGAAATTGTATCATCAACATGAAGGCCGCGAAACTACTACAGT	858
Db	646	CACCCCCTCCTGAGTTTGTACAGAAGCTACAGAGTGTGAGGCCCAAGAACTACATGGAAG	705
Qy	859	CTCTGCGCTCCACAGACCAAGGTGCTTGGGCCAAGCTTTTCCCCAAGTCAGACTCCCAAG	918
Db	706	GCCTCCCTCGATTTTGGAAAGAAGATTTTGTCTCTGTCGACCAATGCAAGGCCCTCAGG	765
Qy	919	CCCTTGACCTGCTCGACCGGATGTTAACTTTTAAACCCCAATAAACGGATCACAGTGGAGG	978
Db	766	CCGTGAATCTCCTGGAAAGAAGATGCTGTGTGTTGGATGCGGAACACGCGGTGCACGACGTG	825
Qy	979	AAGCGCTGGCTCACCCCTACTCTGGAGCAGTACTATGACCCGACGGATGAGCCAGTGGCCG	1038
Db	826	AGGCATTAGCCCAACCCATACTTTTGAGTCCCTTTCCGGACACTTGAGGATGAGCCCAAGGCC	885
Qy	1039	AGGAGCCCTTCACTTTCCGCATGAGCTGGATGACTTACCTTAAGGAGCGCTGAAGCAGC	1098
Db	886	AGAAATATGATGATCTCTTTGATGACGTAGACCCGACCCCTTTGAGGAATGGAAGCGTGTTA	945
Qy	1099	TCATCTTCCAGGAGACAGCAGCTTCCAGCCCGGAGTGCTGGAGGCCCTTCTAGGCCAGAC	1158
Db	946	CGTATAGGAGTGTCTCAGCTTTCAGCTTCCAGCGCAGCTAGGAGCCAGAGTTCCAAAGG	1005
Qy	1159	AGACATCTCTG	1169
Db	1006	AGACAGTCTG	1016

RESUL, T 11

```

RES001.11
US-11-136-527-2977
; Sequence 2977, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2977
; LENGTH: 3824
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2977

```

Query Match	11.1%;	Score	203.8;	DB	8;	Length	3824;
Best Local Similarity	56.5%;	Pred.	N.M.	2.7e-37;			
Matches	420;	Conservative	0;	Mismatches	317;	Indels	6;
Gaps	2;						
QY	336	CGGTCCACCCCTGGAGCCATGAGAGATGCTCTACATTGTGTCAGGACCTCATGTCGAGACTGTAC	395				
DB	672	GTGGGCTCTCTAACAGAGCTGATAGCGCTCTACATTGTTCCAGGATCATGAGACGACAC	731				
QY	396	CTGTACAAGTTGCTGAAAGCCGACGCTGAGCAATGACCCATATCTGTCTATCTTCTCTTAC	455				
DB	732	TTGGCGAACGTCCTGGAGCAGGGCCCTTTATCTGGAGGAGCATGCCAGGCTCTTCTATGTATC	791				
QY	456	CAGATCTCTGGCGGGCTCAAGTATACATCCACTCCGCCAACGTGCTCCACCGAGATCTTAAG	515				
DB	792	CAGCTGCTGCTGGGCTCAAGTATACATCCACTCTGCAACGTCGCTCCACGAGGATCTCAAG	851				
QY	516	CCCTCCAACTGCTCATCAACAC---CACCTGCGACCTTAAGATTTGTGATTTGGGCGCTG	572				
DB	852	CCGGCCCAACCTTTTTCATTAAACACTGACACTGTGGTGTGTAAGATTTGGTGACTTTGGGCTG	911				

QY	573	GCCCGGATTCGGATCTCTGAGCATGACCAACACGGCTTCTCTGACGAGTATGTGGCTACG	632
DB	912	GCCCGGATCATGGATCTCTATTTCCATAAGGTCATCTTTCTGAGGATTGGTTACC	971
QY	633	CGCTGGTACCGGGCCCCAGAGATCATGCTGAACCTCCAAGGCTATATACCAAGTCCATCGAC	692
DB	972	AAATGGTACAGATCTCCACGGCTTTTACTTTCTCTAAATACTATATACTAAGCCATTGAC	1031
QY	693	ATCTGGTCTGTGGCTGCATTCTGGCTGAGATGCTCTCTAAACGGGCCCATCTTCCCTGGC	752
DB	1032	ATGTGGGCTCGAGGCTGCATCTTTGTCTGAATGCTGACTGGTAAACCCCTCTTTGACGGT	1091
QY	753	AAGCACTACTCTGATCAGCTCAACCAATCTCTGGGCATCTCTGGGCTCCCATCCCAAGGAG	812
DB	1092	GCACATGAATTTGAACAGATGCAGCTGATCTTGGAGTCTATCCTCTGTGTGCACGAGAA	1151
QY	813	GACCTGAATGTTATCATCAACATGAGGCGCGAAACTACCTACAGTCTCTGCCCTCCAAG	872
DB	1152	GATCGGCAGGAGCTTCTCAGCGTGAATCCAGTTTACATTTAGAACGACATG---ACTGAG	1208
QY	873	ACCAAGGTGGCTTGGGCGAAGCTTTTCCCAAGTCAGACTCCAAAGCCCTTGACCTGCTG	932
DB	1209	CCACAAAACGCTGACTCAGCTGCTTCCGGGATTAGTCGGGAAGCACTGGATTTCCTG	1268
QY	933	GACCGGATGTTAACTTTAAACCCCAATPAAACGATCACAGTGGAGGAGCGCTGGCTCAC	992
DB	1269	GAAACAGATTCTGACGTTTCAGTCCCATGGACCGGCTGACAGCCGAGGAAGCACATTTCCCAT	1328
QY	993	CCCTTACCTGGAGCAGTACTATGACCCGACGATGAGCCAGTCAGTCGGCCGAGGAGCCCTTCACC	1052
DB	1329	CCTTACATGAGCATCTACTCTTTCCTCCAAACGACGAGGCTTATTCAGCCATCCTTTCCAC	1388
QY	1053	TTCCGCATGGAGCTGGAGTACCT	1075
DB	1389	ATAGAAGACGAAGTGGACGACAT	1411

RESULT 12

```

RESULTS 12
US-11-136-527-3508
; Sequence 3508, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM1010867)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3508
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3508

```

	Query Match	9.1%	Score 166.4	DB 8	Length 822
	Best Local Similarity	56.8%	Pred. No. 6.2e-29		
	Matches 349	Conservative 0	Mismatches 256	Indels 9	Gaps 2
QY	355	TTGAGAGATGTTACATTGTGCAGGACCTGTATGGAGACTGACCTGTACAGTTGCTGAAA	414		
Db	215	TCAGCGTGGCTTATTCTTCCAGAGTACATGGAGACTGACCTGGCATGCTGCTAGAGC	274		
QY	415	GCCAGCAGCTGAGCAATGACCATATCTGCTACTTCTCTACAGATCTCTGGGGGGCTCA	474		
Db	275	AGGGCACGCTGACCGGAGGACCGCTAAGCTATTATGATGACAGCTGCTGGCTGGGCTCA	334		
QY	475	AGTACATCCACTCGCCCAACGTGCTTCCACCGAGATCTTAAAGCCCTCCAACTGCTCTATCA	534		

Db 335 AGTATACCTCACTCGCCAAAGTGTTCACAGGACCTGAAGCCCGCAACATTTTCATCA 394
QY 535 ACAC---CACCTGGACCTTAAGATTGTGATTTTCGGCCCTGGCCCGGATTCGGATCCTG 591
Db 395 GCACGGAGGACCTCGTCTGAAGATCGGGATTTTCGGGCTGGCCAGAAATCGCGGACAGC 454
QY 592 AGCATGACACACCGGCTTCCTGACGAGTATGTCGTACGCTACGCGCTGTAACGGGCCCCAG 651
Db 455 ATTACTCCCAAGGGTTATCTGTGGAAGGGTTGTGTGCAAAAGTGGTACCGCTCTCCAC 514
QY 652 AGATCATGTGAACTCCAAAGGCTATACCAAGTCCATCGACATCTGTGTGTGGGCTGCA 711
Db 515 GACTGTCTGTGCCGCAACCTACACGAAGCCATCGATATGTGGCAGCTGGCTGCA 574
QY 712 TTCGGCTGAGTCTCTTAACCGGCCCCATCTTCCTGTCGCAAGCACTACCTGGATCAGC 771
Db 575 TCCTAGCGGAGATGCTCACGGGAAATGCTCTTGTCTGGGCTCACGAGCTTGAGCAGA 634
QY 772 TCAACACATTTCTGGGCATCTCGGCTCCCATCCAGGAGGACCTGAATTGTATCATCA 831
Db 635 TGAGCTTCATCTTAGACACCATCCCTGTAGTGGGGAGGAGAACAGGAGGAGCTGCTCA 694
QY 832 ACATGAAGGCCGAAACTACTACAGTCTCTGCTCCCTCCAAGACCAAGGTGGCTTGGGCCA 891
Db 695 GGGTGTGCC-----GTCTTTGTGACGAGCACTTGGGAGGTGAAGAGGCCACTGCGTA 748
QY 892 AGCTTTTCCCAAGTCAGACTCAAAGCCCTTGACCTGTGACCGCCGATTTAACCTTTA 951
Db 749 AGCTACTCCCGGATGTCAACCGTGAAGCCATTGACTTCTTGGAGAAGATCCTGACGTTCA 808
QY 952 ACCCCCAATAACGG 965
Db 809 GCCCATGGACCGG 822

RESULT 13
US-11-136-527-238
; Sequence 238, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 238
; LENGTH: 1980
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-238

Query Match 9.0%; Score 165.8; DB 8; Length 1980;
Best Local Similarity 55.7%; Pred. No. 1.1e-28;
Matches 452; Conservative 0; Mismatches 332; Indels 27; Gaps 6;

QY 115 CGTTCGACGTGGGCGCGCTACACGAGTTGCAGTACATCGCGGAGGGCGGTACGGCA 174
Db 419 CTTTCAGTTCTTAAGCGTTCACGAACTTGAAGCGATCGGCTCTGGGCTCAGGAA 478
QY 175 TGCTCAGCTCGGCTATGACCACTGCGCGCAAGACTCGCTGGCCATCAAGAAGATCAGC- 233
Db 479 TAGTTTGTGCTGATGACGCTGTCTCTGACGAAATGTGGCCATTAAAGACTCAGCA 538
QY 234 --CCCTTCGAACATCAGACCTACTGCGAGCGCAGCTCCCGGAGATCCAGATCCTGCTGC 291
Db 539 GACCCCTTCCAGAACCAACTCATGCCAAGAGGGCTTACCGGAGCTGGTCTCATGAAGT 598
QY 292 GCTTCCGCCATGAGAATGTATCGGCATCCGA---GACATTTCTCGCGGCGCTCCACCCCTGG 348

Db 599 GTGTGAACCAATAAACAATTAATTAAGCTTATTAAGTCTTTTACACCCAGAAAAACACTGG 658
QY 349 AAGCCATGAGAGATGTCTACATTGTGACGAGCTGATGAGACTGACCTGTACAAAGTTGC 408
Db 659 AGGAGTTCCAAGATGTTTACTTTAGTATGGAACCTGATGGACGCCAACTTGTGTCAGGTGA 718
QY 409 TGAAGAAGCCAGCAGCTGAGCAATGACCATATCTGCTACTTCTCTACCAAGATCCTTCGGG 468
Db 719 TTCAGA---TGAGCTGGACCAAGCGGATGTCGTACTTGTCTACCAAGATGCTGTGG 775
QY 469 GCCTCAAGTACATCACTCCGCAACGCTCTCCACCGAGATCTAAAGCCCTCCAACTGTC 528
Db 776 CGATCAACACCTCCACTCCGCTGGGATCATCCACAGGACTTAAAAACCCAGTAACATCG 835
QY 529 TCATCAACACCACTGCGACCTTAAGATTGTGATTTGGCTCTGGCCCGGATTCGGGATC 588
Db 836 TAGTCAAGTCTGATTGCACTCAAAATCTCTGGACTTTGGACTGGCCAGGACAGC----- 890
QY 589 CTGAGCATGACACACCGGCTTCTCTGAGGAGTATGTGGCTACGCGCTGGTACCGGGCC 648
Db 891 -----GGGCAAGCTTCATGATGATCCCGTATGTGTGACGAGATATTACAGAGCC 943
QY 649 CAGAGATCATGCTGAACCTCAAAGGCTATPACCAAGTCCATCGACATCTGCTGTGGGCT 708
Db 944 CCGAGGTCACTCG---GGCATGGGCTACAAGAGAGAACGTGGACATATGCTCTGTGGGCT 1000
QY 709 GCATTCTGGCTGAGATGCTCTTAACCGGCCCATCTTCCCTGGCAAGCACTACCTGGATC 768
Db 1001 GCATCATGGGAGAAATGGTTCGTCAAAATCTCTTCCCGGAAGGAGACTATATTGACC 1060
QY 769 AGCTCAACACCATTTCTGGGCATCTCTGGGCTCCCATCCAGGAGGACCTGAATTGTATCA 828
Db 1061 AGTGAACCAAGTCAATAGAGAGCTAGGAATCTCCGTGTCCAGATTCATGAA---GAAAT 1117
QY 829 TCAACATGAAGGCCCGAAACTACTACAGTCTCTGCCCTTCCAGAACCAAGGTGGCTGG 888
Db 1118 TGCAGCCACCGTCAGAACTAGTGGAGAACCGGCCCAAGTATGAGGCGCTCACCTTCC 1177
QY 889 CCAAGCTTTTCCCAAGTCAAGCTCAAAAGC 919
Db 1178 CCAAGCTTTTCCAGATTCCCTCTTCCACG 1208

RESULT 14
US-11-136-527-7604
; Sequence 7604, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7604
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7604

Query Match 9.0%; Score 165; DB 8; Length 600;
Best Local Similarity 57.0%; Pred. No. 1.2e-28;
Matches 344; Conservative 0; Mismatches 250; Indels 9; Gaps 2;

QY 366 TACATTGTGCAGACCTGATGGAGCTGACCTGTACAAAGTTCTGTAAGAGCCAGCAGCTG 425
Db 4 TATATTGTCCAAGAGTACATGAGAGTACCTGCGCATGCTCTAGAGCGGACGCTG 63

426	Qy	ACCAATGACCATATCTGCTACTTCTCTACAGATCTCGGGGCTCTCAAGTACATCCAC	488
64	Db	ACCGAGGAGCAGCTAAGCTATTCATGTACAGCTGCTGGTGGGCTCAAGTACATCCAC	123
486	Qy	TCGCGCAACGTGCTCCACCGAGATCTAAAGCCCTCCAACTGCTCATCAACAC--CAC	542
124	Db	TTCTGCCACGTGTTGCCACAGGACCTGNAAGCCCGCCAAATTTTCATCAGCAGGGAGC	183
543	Qy	TCGCACCTTAAGATTTGTGATTTTCGGGCTTGCCCGGATTTGCCGATCTTGAGCATGACAC	602
184	Db	CTCGTCTGAGATCGGGGATTTTCGGGCTGGCCAGAAATCGCGACAGCAATTAATCCAC	243
603	Qy	ACGGGCTTCGACGGAGTATGTGGCTACGGCTGCTACCGGGCCCGAGAGATCATGCTG	662
244	Db	TAAGGGTTATCTGTTCGGAGGGTTGGTGACAAAGTGGTACCGCTCTCCACAGCTGCTCTG	303
663	Qy	AACTCCAGGGGTATACCAAGTCCATCGACATCTGGTCTGTGGGCTGCATTTCTGGCTGAG	722
304	Db	TCCCGGACCACTACACGAAAGCCATCGATATGTGGGCACTGGCTGCATCCTTAGCGGAG	363
723	Qy	ATGCTCTCTTAACCGGCCCATCTTCCTGGCAAGCACTTACCTGGATCAGCTCAACCAATT	782
364	Db	ATGCTCACGGGAAAATGCTCTTTGCTGGGCTCACGAGCTTGAGCAGATGCAAGTCATC	423
783	Qy	CTGGGATCTCGGGTCCCGATCCAGAGGAGCACTGAAATTGTATCATCAATGAGGCC	842
424	Db	CTAGACACCATCCCTGTAGTCGGGAGGAGAACAGGAGGAGCTGTCTCAGGGTGATGCC-	482
843	Qy	CGAACTACCTACATCTCTGCCCCCTCCAAAGCAAGGTGGCTGGGCAAGCTTTTCCCC	902
483	Db	-----GTCCTTTGTTCAGCAGCACTTGGGAGGTGAAGAGGCCACTGCGTAGCTACTCCCG	537
903	Qy	AAGTCAGACTCCAAAGCCCTTGACCTGCTGGACCGGATGTTTAACCTTTAAACCCCAATAA	962
538	Db	GATGTCAACCGTAGACCATTGACTTTTCTGGAGAGATCCTGAGCTTCAGGCCCAATGGAC	597
963	Qy	CGG 965	
598	Db	CGG 600	

```

RESULT 15
US-11-136-527-4380
; Sequence 4380, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60,574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4380
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (587)..(587)
; OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-4380

Query Match      8.9%; Score 164.2; DB 8; Length 600;
Best Local Similarity 65.2%; Pred. No. 1.8e-28;
Matches 362; Conservative 6; Mismatches 164; Indels 23; Gaps 8;

Qy      1284  ACTGTGCCAGCCCGGACCTTGGCAGCCCGCGGGTGGAGCATGGGCTGGCCACCT 1343
Db      10    ACTTTGCTCAACCGGACCCCGGCGAGCCCAAG-CTGCACCAAGGGTGGGCTGGCACCCC 68

```

[illegible]

Search completed: February 7, 2006, 14:11:06
Job time : 796.931 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:19:31 ; Search time 266.938 Seconds
(without alignments)
551.408 Million cell updates/sec

US-10-623-108-2

Title: Perfect score: 1763

Sequence: 1 MAAAAAQQGGGGPRRTG.....LKELIFQTARFQGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq 21.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*
- 9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	335	8	ADH48364 Human KPP
2	1763	100.0	335	8	ABM82550 Human dia
3	1763	100.0	335	8	ABM82548 Human dia
4	1763	100.0	335	9	ADW12901 Human SMA
5	1731	98.2	379	2	AAW15506 Mitogen a
6	1731	98.2	379	4	AAM38754 Human pol
7	1731	98.2	379	6	ABR43898 Human ERK
8	1731	98.2	379	7	ADH59632 Erkl prot
9	1731	98.2	379	7	ADH59632 Erkl prot
10	1731	98.2	379	8	ADO2522 Biochemic
11	1731	98.2	379	8	ADQ37846 Human pho
12	1731	98.2	379	9	ADV42070 Human ext
13	1731	98.2	379	9	ADW12905 Human SMA
14	1731	98.2	379	9	ADZ65036 Mitogen a
15	1731	98.2	403	4	AAM40540 Human pol
16	1725	97.8	379	2	AAI42413 Extracell
17	1725	97.8	379	4	AAG67618 Amino aci
18	1725	97.8	379	4	AAG67439 Human ext
19	1725	97.8	379	5	ABG31847 Human ext
20	1725	97.8	379	6	ABU89742 Protein d
21	1725	97.8	379	7	ABR62625 Human pro
22	1725	97.8	379	8	ADO24425 Human PRO
23	1725	97.8	379	9	ADZ10062 Human bre
24	1725	97.8	631	2	AAW85006 Erkl-gree

25	1725	97.8	631	3	AAV70778	AAV70778 EGFP-Erkl
26	1659	94.1	624	2	AAW85015	AAW85015 Erkl-gree
27	1649	93.5	375	6	ABP96053	ABP96053 Human pro
28	1607	91.2	356	2	AAW48013	AAW48013 Extracell
29	1606	91.1	355	7	ADC64458	ADC64458 Human ext
30	1606	91.1	355	8	ADG98194	ADG98194 Human ext
31	1599	90.7	359	9	ADW12903	ADW12903 Human SMA
32	1599	90.7	359	9	ADW12907	ADW12907 Human SMA
33	1595	90.5	365	2	AAW20103	AAW20103 ERKL. 8/2
34	1593	90.4	359	6	ABU54596	ABU54596 Human NOV
35	1486.5	84.3	340	8	ABM82549	ABM82549 Human dia
36	1486.5	84.3	340	8	ABM82547	ABM82547 Human dia
37	1454.5	82.5	380	7	ADC64459	ADC64459 Human ext
38	1454.5	82.5	380	8	ADG98195	ADG98195 Human ext
39	1444.5	81.9	360	5	ABG31848	ABG31848 Human cer
40	1444.5	81.9	360	6	ABR43899	ABR43899 Human ERK
41	1444.5	81.9	360	7	ADH59634	ADH59634 Erkl2 prot
42	1444.5	81.9	360	7	ADH59634	ADH59634 Erkl2 prot
43	1444.5	81.9	360	8	ADJ66608	ADJ66608 ERK prote
44	1444.5	81.9	360	8	ADJ66608	ADJ66608 ERK prote
45	1444.5	81.9	360	8	ADP56268	ADP56268 Human PRO

ALIGNMENTS

RESULT 1
ADH48364
ID ADH48364 standard; protein; 335 AA.
XX ADH48364;
XX XX
DT 25-MAR-2004 (first entry)
XX XX
DE Human KPP protein SEQ ID NO:22.
XX human; KPP; kinase; phosphatase; enzyme; cytostatic; anorectic;
KW immunosuppressive; KPP-Antagonist; KPP-Agonist; gene therapy;
KW autoimmune disorder; obesity; cancer.
XX Homo sapiens.
XX XX
PN WO2004001008-A2.
XX XX
PD 31-DEC-2003.
XX XX
PF 19-JUN-2003; 2003WO-US019660.
XX XX
PR 21-JUN-2002; 2002US-0390652P.
PR 15-JUL-2002; 2002US-0396196P.
(INCY-) INCYTE CORP.
Marquis JP, Baughn MR, Tran UK, Hafalia AJA, Kable AE;
Emerling BM, Elliott VS, Lindquist EA, Richardson TW, Khare R;
Swarnakar A, Lee SY, Ramkumar J, Chawla NK, Becha SD, Mason PM;
Hawkins PR, Bulloch SA, Jin P, Bhatia U, Burrill JD, Lee S;
Blake JJ, Ho A, Zheng W;
WPI; 2004-082489/08.
N-PSDB; ADH48413.
New human kinases and phosphatases (KPP) polypeptide, useful for
preparing a composition for treating a disease associated with decreased
expression or overexpression of functional KPP e.g., cancer.
Claim 1; SEQ ID NO 22; 336pp; English.
The invention relates to novel isolated human kinases and phosphatases
(KPP) polypeptides. A protein of the invention has cytostatic, anorectic,
and immunosuppressive activity, and acts as a KPP-Antagonist, or KPP-
Agonist. A polynucleotide of the invention may have a use in gene
therapy. The polypeptide is useful for preparing a composition for

CC	diagnosing or treating a disease or condition associated with decreased expression or overexpression of functional KPP e.g., autoimmune disorders, obesity or cancer. The sequences shown in ADH48343-ADH48391 represent KPP proteins of the invention.	PT	New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.
XX		XX	
SQ	Sequence 335 AA;	PS	Claim 27; Page; 190pp; English.
		CC	The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorders, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp protein of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
		XX	Sequence 335 AA;
		SQ	
			Query Match 100.0%; Score 1763; DB 8; Length 335;
			Best Local Similarity 100.0%; Pred. No. 5.2e-174; Indels 0; Gaps 0;
			Matches 335; Conservative 0; Mismatches 0;
QY	1 MAAAAQGGGGPRTEGVPGEVEMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60	QY	1 MAAAAQGGGGPRTEGVPGEVEMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
DB	1 MAAAAQGGGGPRTEGVPGEVEMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60	DB	1 MAAAAQGGGGPRTEGVPGEVEMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
QY	61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFHENVIGIRDLRASTLEAMRDVYI 120	QY	61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFHENVIGIRDLRASTLEAMRDVYI 120
DB	61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFHENVIGIRDLRASTLEAMRDVYI 120	DB	61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFHENVIGIRDLRASTLEAMRDVYI 120
QY	121 VQDLMETDLYKLKSKQSLNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180	QY	121 VQDLMETDLYKLKSKQSLNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB	121 VQDLMETDLYKLKSKQSLNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180	DB	121 VQDLMETDLYKLKSKQSLNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
QY	181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240	QY	181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB	181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240	DB	181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY	241 NRPIFGKHLYDQNLHILALDLDRLMTPNPNKRITVEEALAHPLYEQYYDPTDEPVAEE 300	QY	241 NRPIFGKHLYDQNLHILALDLDRLMTPNPNKRITVEEALAHPLYEQYYDPTDEPVAEE 300
DB	241 NRPIFGKHLYDQNLHILALDLDRLMTPNPNKRITVEEALAHPLYEQYYDPTDEPVAEE 300	DB	241 NRPIFGKHLYDQNLHILALDLDRLMTPNPNKRITVEEALAHPLYEQYYDPTDEPVAEE 300
QY	301 PFTFAMELDDLPKRLKELIFOETARFQGVLEAP 335	QY	301 PFTFAMELDDLPKRLKELIFOETARFQGVLEAP 335
DB	301 PFTFAMELDDLPKRLKELIFOETARFQGVLEAP 335	DB	301 PFTFAMELDDLPKRLKELIFOETARFQGVLEAP 335
			RESULT 3
			ABM82548
ID	ABM82550 standard; protein; 335 AA.	ID	ABM82548 standard; protein; 335 AA.
XX		XX	
AC	ABM82550;	AC	ABM82548;
XX		XX	
DT	18-NOV-2004 (first entry)	DT	18-NOV-2004 (first entry)
XX		XX	
DE	Human diagnostic and therapeutic pprotein SEQ ID NO:2799.	DE	Human diagnostic and therapeutic pprotein SEQ ID NO:2797.
XX		XX	
KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX		XX	
OS	Homo sapiens.	OS	Homo sapiens.
XX		XX	
PN	WO2004023973-A2.	PN	WO2004023973-A2.
XX		XX	
PD	25-MAR-2004.	PD	25-MAR-2004.
XX		XX	
PF	12-SEP-2003; 2003WO-US028227.	PF	
XX		XX	
PR	12-SEP-2002; 2002US-0410259P.	PR	
XX		XX	
PR	12-SEP-2002; 2002US-0410260P.	PR	
XX		XX	
PA	(INCY-) INCYTE CORP.	PA	
XX		XX	
PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;	PI	
PI	Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Eider LV;	PI	
PI	Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;	PI	
PI	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;	PI	
PI	Paraita CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;	PI	
PI	Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;	PI	
PI	Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;	PI	
PI	Patury S, Shi X, Suarez CJ;	PI	
XX		XX	
DR	WPI; 2004-329368/30.	DR	
DR	N-PSDB; ACN41202.	DR	
XX		XX	

```

PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtan ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR WPI; 2004-329368/30.
DR N-PSDB; ACN41200.
XX
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 335 AA;
Query Match 100.0%; Score 1763; DB 8; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.2e-174;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAAAAGGGGGPRTEGVPGEVEMVKQPFDPVGRYTQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRTEGVPGEVEMVKQPFDPVGRYTQLOVIGEGAYGMVSSAY 60
Qy 61 DHVKTRVAIKKISPFPHQTYCQRTLEIRIQLLRFHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVKTRVAIKKISPFPHQTYCQRTLEIRIQLLRFHENVIGIRDILRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYTHSANVLHRLDKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYTHSANVLHRLDKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
Qy 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
Db 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
Qy 301 PFTFAMELDDLKPKERLKELIQFQETARFQGVLEAP 335
Db 301 PFTFAMELDDLKPKERLKELIQFQETARFQGVLEAP 335

RESULT 4
ADW12901
ID ADW12901 standard; protein; 335 AA.
XX
AC ADW12901;
XX
DT 07-APR-2005 (first entry)
DE Human SMAPK3V1 variant protein.
XX
KW DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW enzyme.
XX
OS Homo sapiens.
XX
PN US2005013817-A1.
XX
PD 20-JAN-2005.
XX
PF 18-JUL-2003; 2003US-00623108.
XX
PR 18-JUL-2003; 2003US-00623108.
XX
PA (DAIK/) DAI K.
XX
PI Dai K;
XX
DR WPI; 2005-080923/09.
DR N-PSDB; ADW12900.
XX
XX New isolated SMAPK3 polypeptides and nucleic acids, useful for diagnosing
PT diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene
PT in a mammal.
XX
PS Claim 1; SEQ ID NO 2; 55pp; English.
XX
XX The invention relates to novel isolated polypeptide (I) comprising an
CC amino acid sequence selected from sequences comprising 335 or 359 amino
CC acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
CC polynucleotides are useful for diagnosing diseases, e.g. cancers,
CC associated with the deficiency of the SMAPK3 gene in a mammal. The
CC fragments of the polypeptides and polynucleotides can also be used as
CC primers or probes. This sequence corresponds to the SMAPK3V1 variant
CC protein.
XX
SQ Sequence 335 AA;
Query Match 100.0%; Score 1763; DB 9; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.2e-174;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAAAAGGGGGPRTEGVPGEVEMVKQPFDPVGRYTQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRTEGVPGEVEMVKQPFDPVGRYTQLOVIGEGAYGMVSSAY 60
Qy 61 DHVKTRVAIKKISPFPHQTYCQRTLEIRIQLLRFHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVKTRVAIKKISPFPHQTYCQRTLEIRIQLLRFHENVIGIRDILRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYTHSANVLHRLDKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYTHSANVLHRLDKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
Qy 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
Db 241 NRPIFPKGHYLDQNLHIALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEE 300
Qy 301 PFTFAMELDDLKPKERLKELIQFQETARFQGVLEAP 335
Db 301 PFTFAMELDDLKPKERLKELIQFQETARFQGVLEAP 335
```

|||||
301 PTFANELDDLPERLKELIFQETARFQGVLEAP 335

Db	301 PTFANELDDLPERLKELIFQETARFQGVLEAP 335	QY	1 MAAAAQGGGGEPRRTTEGVGVGEVEMVKGQPFVGPRTYQIQYIGEGAYGMVSSAY 60
XX		Db	1 MAAAAQGGGGEPRRTTEGVGVGEVEMVKGQPFVGPRTYQIQYIGEGAYGMVSSAY 60
XX		QY	61 DHVTRTRVAIKKISPFHQYTCQRTLREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
XX		Db	61 DHVTRTRVAIKKISPFHQYTCQRTLREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
XX		QY	121 VQDLAMETDLYKLLKSQQLSNDHICYFLYQILLRGLKYIHSANVLRDLKPSNLLINTCDL 180
XX		Db	121 VQDLAMETDLYKLLKSQQLSNDHICYFLYQILLRGLKYIHSANVLRDLKPSNLLINTCDL 180
XX		QY	181 KICDFGLARIADPEHDHTGFTTEYVATRYRAPEIMLNSKGYTKSIDISWVCILAEMLS 240
XX		Db	181 KICDFGLARIADPEHDHTGFTTEYVATRYRAPEIMLNSKGYTKSIDISWVCILAEMLS 240
XX		QY	241 NRPIFGKHYLDQLNHIL----- 258
XX		Db	241 NRPIFGKHYLDQLNHIL----- 258
XX		QY	259 --ALDLLDRMLTFNPKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPERKEL 316
XX		Db	301 SKALDLLDRMLTFNPKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPERKEL 360
XX		QY	317 KELIFQETARFQGVLEAP 335
XX		Db	361 KELIFQETARFQGVLEAP 379
XX		RESULT 6	
XX		AAW38754	
XX		ID	AAW38754 standard; peptide; 379 AA.
XX		AC	AAW38754;
XX		DT	22-OCT-2001 (first entry)
XX		DE	Human polypeptide SEQ ID NO 1899.
XX		KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX		KW	peripheral nervous system; neuropathy; central nervous system; CNS;
XX		KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX		KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX		KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX		KW	leukaemia.
XX		OS	Homo sapiens.
XX		PN	WO200153312-A1.
XX		PD	26-JUL-2001.
XX		PF	26-DEC-2000; 2000WO-US034263.
XX		PR	23-DEC-1999; 99US-00471275.
XX		PR	21-JAN-2000; 2000US-00488725.
XX		PR	25-APR-2000; 2000US-00552317.
XX		PR	20-JUN-2000; 2000US-00598042.
XX		PR	19-JUL-2000; 2000US-00620312.
XX		PR	03-AUG-2000; 2000US-00653450.
XX		PR	14-SEP-2000; 2000US-00662191.
XX		PR	19-OCT-2000; 2000US-00693036.
XX		PR	29-NOV-2000; 2000US-00727344.
XX		PA	(HYSE-) HYSEQ INC.
XX		PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX		PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX		PI	Zhou P, Goodrich R, Drmanac RT;
XX		DR	WPI; 2001-442253/47.
XX		DR	N-PSDB; AAI57910.

RESULT 5
AAW15506
ID AAW15506 standard; peptide; 379 AA.
AC AAW15506;
DT 03-JUN-1997 (first entry)
DE Mitogen activating protein kinase ERK1.
XX Antibody; MAP kinase; human; mitogen activated protein; lymphoid cell;
KW recognition sequence; IgG; immunogen; rat; ERK2; hydridoma; spleen cell;
KW MAP kinase-related disease; cancer-associated disease; brain tumour;
KW metabolic disorder; diabetes mellitus; circulatory disease; rheumatism;
KW arteriosclerosis; allergic diseases; central nervous system disease;
KW asthma; Alzheimer's disease; Parkinson's disease; senile dementia; ERK1;
KW bone/joint disease; pollenosis; atopic dermatitis.
XX Homo sapiens.
XX EP735370-A1.
XX 02-OCT-1996.
XX 28-MAR-1996; 96EP-00104966.
XX 28-MAR-1995; 95JP-00070125.
XX 24-NOV-1995; 95JP-00305456.
XX 08-DEC-1995; 95JP-00320577.
XX (TAKE) TAKEDA CHEM IND LTD.
XX Tani A, Ichimori Y;
XX WPI; 1996-444623/45.
XX New antibodies to human MAP kinase - used for the detection, assay and
XX purification of various MAP kinase species.
XX Example 3; Page 34-35; 59pp; English.
XX This sequence represents the human mitogen activated protein (MAP) kinase
XX ERK1. Fragments of this sequence (such as AAW15501) are recognition
XX sequences for the antibodies of the invention. The antibodies of the
XX invention are IgG type monoclonal antibodies (Mab) generated by using
XX human ERK1 as an immunogen and are capable of binding to ERK1 at the
XX sites represented by these peptides. However, the Mab is incapable of
XX binding to human or rat MAP kinase ERK2, or rat MAP kinase ERK1. The Mab
XX are produced from cloned hydridomas derived from spleen cells of a mammal
XX immunised with human ERK1 and homologous or heterologous lymphoid cells.
XX The antibodies can be used for detecting or assaying MAP kinases. They
XX can be used for diagnosing MAP kinase-related diseases, such as cancer-
XX associated diseases (e.g. brain tumour), metabolic disorders (e.g.
XX diabetes mellitus), circulatory disease (e.g. arteriosclerosis), allergic
XX diseases (e.g. asthma, pollenosis, atopic dermatitis), central nervous
XX system diseases (e.g. Alzheimer's disease, Parkinsonism, senile dementia)
XX and bone/joint diseases (e.g. rheumatism). They can also be used to
XX investigate the role of MAP kinases in-vivo and the mechanism of drug
XX action. The antibodies can also be used for purifying MAP kinases. By
XX using antibodies specific for various types of MAP kinase, the activity
XX of any given species of MAP kinase can be detected or assayed with high
XX sensitivity to the exclusion of other species of MAP kinase
XX Sequence 379 AA;
XX Query Match 98.2%; Score 1731; DB 2; Length 379;
XX Best Local Similarity 88.4%; Pred. No. 1.3e-170;
XX Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.

XX Example 3; SEQ ID NO 1899; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 4; Length 379;

Best Local Similarity 88.4%; Pred. No. 1.3e-170;

Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPFVDPGYTQIQYIGSGAYGMVSSAY 60

DB 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPFVDPGYTQIQYIGSGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPFPHQYTCQRTLEIQILFRHENVIGIRDLASTLEAMRDVYI 120

DB 61 DHVRKTRVAIKKISPFPHQYTCQRTLEIQILFRHENVIGIRDLASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180

DB 121 VQDLMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

DB 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

QY 241 NRPIFGKHYLDQNLHIL----- 258

DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPKSD 300

QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFAMELDDLPKERL 316

DB 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFAMELDDLPKERL 360

QY 317 KELIFQETARFQPGVLEAP 335

DB 361 KELIFQETARFQPGVLEAP 379

RESULT 7
ABR43898

ID ABR43898 standard; protein; 379 AA.

XX AC ABR43898;

XX 11-AUG-2003 (first entry)

XX Human ERK1 protein.

XX DARPP-32; psychotic disorder; intracellular signaling protein;
KW dopamine- and cAMP-regulated phosphoprotein; ERK1; ERK2; CREB;
KW extracellular signal-regulated protein kinase; phosphorylation;
KW cAMP-response element binding protein; schizophrenia; human.

XX Homo sapiens.

XX

PN WO2003021225-A2.

PD 13-MAR-2003.

XX 03-SEP-2002; 2002WO-US027802.

XX 31-AUG-2001; 2001US-0316338P.

XX (UYRQ) UNIV ROCKEFELLER.

PA (KARO-) KAROLINSKA INST.

XX Greengard P, Pisone G;

XX WPI; 2003-300912/29.

XX Identifying agent to be tested for ability to treat psychotic disorder,
PT by contacting cells/tissues with candidate drug, determining
PT phosphorylation levels of intracellular signaling proteins DARPP-32,
PT ERK1, ERK2, CREB.

XX Disclosure; Page 77-78; 79pp; English.

XX The invention relates to identifying an agent to be tested for ability to
CC treat psychotic disorder in patient. The method involves contacting cells
CC or tissues with a candidate drug, and determining levels of
CC phosphorylation of intracellular signaling proteins DARPP-32 (dopamine-
CC and cAMP-regulated phosphoprotein), ERK1 and ERK2 (extracellular signal-
CC regulated protein kinases 1 and 2), and CREB (cAMP-response element
CC binding protein). The method is useful for identifying an agent to be
CC tested for an ability to treat a psychotic disorder such as schizophrenia
CC in a patient in need of such treatment. The present sequence represents a
CC human ERK1 protein

XX Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 6; Length 379;

Best Local Similarity 88.4%; Pred. No. 1.3e-170;

Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPFVDPGYTQIQYIGSGAYGMVSSAY 60

DB 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGQPFVDPGYTQIQYIGSGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPFPHQYTCQRTLEIQILFRHENVIGIRDLASTLEAMRDVYI 120

DB 61 DHVRKTRVAIKKISPFPHQYTCQRTLEIQILFRHENVIGIRDLASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180

DB 121 VQDLMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

DB 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

QY 241 NRPIFGKHYLDQNLHIL----- 258

DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPKSD 300

QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFAMELDDLPKERL 316

DB 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFAMELDDLPKERL 360

QY 317 KELIFQETARFQPGVLEAP 335

DB 361 KELIFQETARFQPGVLEAP 379

RESULT 8

ADF45050

ID ADF45050 standard; protein; 379 AA.

XX AC ADF45050;

XX 12-FEB-2004 (first entry)
DT Human kinase ERK1.
XX Human; protein kinase; enzyme; inhibitor; ERK1.
XX Homo sapiens.
XX WO2003081210-A2.
PN 02-OCT-2003.
XX 20-MAR-2003; 2003WO-US008725.
XX 21-MAR-2002; 2002US-0366892P.
XX (SUNE-) SUNESIS PHARM INC.
PA Prescott JC, Braisted A;
PI WPI; 2003-865136/80.
XX Identifying ligand binding to inactive conformation of target protein
XX kinase (T) comprises contacting the conformation modified (T) which
XX contains reactive group at binding site, with ligands and detecting
XX kinase-ligand conjugate formation.
XX Disclosure; SEQ ID NO 19; 260pp; English.
XX The present invention relates to a method for identifying a ligand (L),
XX which binds to an inactive conformation of target protein kinase (T). The
XX method involves contacting inactive conformation of (T), which contains
XX or is modified to contain a reactive group at or near a binding site of
XX interest, with one or more ligand candidates capable of covalently
XX bonding to the reactive group thus forming a kinase-(L) conjugate (C).
XX The method is useful for identifying protein kinase inhibitors that
XX preferentially bind to inactive conformation of a target protein kinase.
XX The present sequence is a protein kinase which may be modified via an
XX amino acid substitution, for use in the method of the invention.
XX SQ Sequence 379 AA;
Query Match 98.2%; Score 1731; DB 7; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAQQGGGGPRRTGEGVPGVGVEMVKGQPDFVGPRTYQLOYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGPRRTGEGVPGVGVEMVKGQPDFVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVTRKTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
DB 61 DHVTRKTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLRIADPEHDHTGFLTYVATRWYRAPEIMLNSKGYTKSIDIVSGVGLIAEMLS 240
DB 181 KICDFGLRIADPEHDHTGFLTYVATRWYRAPEIMLNSKGYTKSIDIVSGVGLIAEMLS 240
QY 241 NRPIFFGKHVLDQLNHIL----- 258
DB 241 NRPIFFGKHVLDQLNHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAMAKLPKSD 300
QY 259 --ALDLDRLMTNPNKRIITVEALAHPLYEQYDPTDEPVAEPTFAMELDLPERL 316
DB 301 SKALDLDRLMTNPNKRIITVEALAHPLYEQYDPTDEPVAEPTFAMELDLPERL 360
QY 317 KELIFQETARFQGVLEAP 335
|||||

DB 361 KELIFQETARFQGVLEAP 379
RESULT 9
ADH59632
ID ADH59632 standard; protein; 379 AA.
XX ADH59632;
XX 25-MAR-2004 (first entry)
XX Erk1 protein.
XX cardiac disease; Raf-1; MEK1; Cardiant; telethonin; ϵ agr-actinin;
XX hypertrophy; MEK2; Erk1; Erk2.
XX Homo sapiens.
XX WO2003025205-A2.
XX 27-MAR-2003.
XX 18-SEP-2002; 2002WO-EP010489.
XX 19-SEP-2001; 2001US-0323566P.
XX 24-SEP-2001; 2001US-0324625P.
XX (MEDI-) MEDIGENE AG.
XX Nave B, Roenicke V, Leclair S, Funk M, Reuner B, Brinkmann K;
XX Henkel T;
XX WPI; 2003-371821/35.
XX N-PSDB; ADH59631.
XX Identifying and/or obtaining a compound useful for preventing or treating
XX cardiac diseases, particularly congestive heart failure, comprises
XX quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
XX on the compound.
XX Claim 33; SEQ ID NO 8; 41pp; English.
XX The present invention relates to identifying and obtaining a compound
XX useful in the prevention or treatment of cardiac diseases, comprising
XX quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
XX on the compound. The method is useful for identifying and/or obtaining
XX compounds that may be used in the prevention or treatment of cardiac
XX diseases, particularly congestive heart failure. The compound or protein
XX is also used for the preparation of a pharmaceutical composition for
XX prevention or treatment of a disease related to hypertrophy or impaired
XX or increased activation of telethonin (T-Cap), ϵ agr-actinin, MHC, actin,
XX titin, myomesin, nebulin, tropomyosin, troponin, Erk1/2 and/or MLCK.
XX The protein or antibody which specifically recognizes the
XX activated/phosphorylated form of the above polypeptide, is used for the
XX preparation of a composition for diagnosing a disease or a predisposition
XX for a disease related to hypertrophy or related to impaired or increased
XX activation of Raf-1, MEK1/2 and/or Erk1/2. The present sequence
XX represents Erk1.
SQ Sequence 379 AA;
Query Match 98.2%; Score 1731; DB 7; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAQQGGGGPRRTGEGVPGVGVEMVKGQPDFVGPRTYQLOYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGPRRTGEGVPGVGVEMVKGQPDFVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVTRKTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
DB 61 DHVTRKTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120

Qy 121 VQDLMTDLYKLLKSQLSNDHICFYQLIRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMTDLYKLLKSQLSNDHICFYQLIRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240
Qy 241 NRPIFGKHYLDQNLHIL----- 258
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLPPKSD 300
Qy 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 316
Db 301 SKALLDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 360
Qy 317 KELIFOETARFQPGVLEAP 335
Db 361 KELIFOETARFQPGVLEAP 379

RESULT 10

ID ADO22522 standard; protein; 379 AA.

XX ADO22522;

DT 12-AUG-2004 (first entry)

DE Biochemical pathway-related human ERK protein SeqID8.

XX biochemical pathway; mammalian glioma tumour;
KW epidermal growth factor receptor; EGFR; mTOR; polypeptide inhibitor;
KW cancer; ERK; human.

XX Homo sapiens.

XX WO2004044218-A2.

XX 27-MAY-2004.

XX 05-NOV-2003; 2003WO-US035115.

XX 05-NOV-2002; 2002US-0423777P.

XX (REGC) UNIV CALIFORNIA.

PA (CELL-) CELL SIGNALING TECHNOLOGY INC.

XX Mischel PS, Sawyers CL, Smith BL, Crosby K;

XX WPI; 2004-411736/38.

XX Use of biochemical pathways associated with glioblastoma for, e.g.
PT identifying a mammalian glioma tumor that is likely to respond to an
PT epidermal growth factor receptor polypeptide inhibitor or an mTOR
PT polypeptide inhibitor.

XX Claim 1; SEQ ID NO 8; 66pp; English.

XX This invention relates to a novel use of biochemical pathways for
CC identifying a mammalian glioma tumor that is likely to respond to an
CC epidermal growth factor receptor (EGFR) polypeptide inhibitor or an mTOR
CC polypeptide inhibitor, or identifying a mammalian glioma tumor that does
CC not express or expresses a PTEN polypeptide and which is likely to
CC respond or not likely to respond to an inhibitor of mTOR polypeptide
CC activity or inhibitor of EGFR polypeptide activity, respectively. The
CC biochemical pathways are, in particular, disregulated in pathologies such
CC as cancer. The present sequence is that of the human ERK protein which
CC may be used in the method of the invention.

XX Sequence 379 AA;

XX Query Match 98.2%; Score 1731; DB 8; Length 379;

Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

Qy 1 MAAAAAQQGGGGPRTEGVGPGEVEMVKGQPPDVGPRTYQLQYIGEGAYGMVSSAY 60

Db 1 MAAAAAQQGGGGPRTEGVGPGEVEMVKGQPPDVGPRTYQLQYIGEGAYGMVSSAY 60

Qy 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120

Db 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120

Qy 121 VQDLMTDLYKLLKSQLSNDHICFYQLIRGLKYIHSANVLRDLKPSNLLINTTCDL 180

Db 121 VQDLMTDLYKLLKSQLSNDHICFYQLIRGLKYIHSANVLRDLKPSNLLINTTCDL 180

Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240

Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLS 240

Qy 241 NRPIFGKHYLDQNLHIL----- 258

Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLPPKSD 300

Qy 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 316

Db 301 SKALLDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 360

Qy 317 KELIFOETARFQPGVLEAP 335

Db 361 KELIFOETARFQPGVLEAP 379

RESULT 11

ADOQ37846

ID ADOQ37846 standard; protein; 379 AA.

XX ADOQ37846;

DT 09-SEP-2004 (first entry)

DE Human phosphorylated ERK protein, target for HER-2 therapy SeqID 3.

XX human; HER-2 directed therapy; tumour; insulin growth factor receptor;
KW EGFR; epidermal growth factor receptor; EGFR; S6 ribosomal protein; AKT;
KW NDF; ERK; cancer therapy; predictive biomarker; HER-2/neu.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 202 /note= "Optionally a phosphorylated serine (if threonine
FT at 204 is unmodified)"

FT Modified-site 204 /note= "Optionally a phosphorylated threonine (if serine
FT at 202 is unmodified)"

XX WO2004053497-A2.

XX 24-JUN-2004.

XX 11-DEC-2003; 2003WO-US039770.

XX 11-DEC-2002; 2002US-0432942P.

XX (VENT-) VENTANA MEDICAL SYSTEMS INC.
XX (CELL-) CELL SIGNALING TECHNOLOGY INC.

XX Bacus SS, Smith BL;

XX WPI; 2004-507069/48.

XX Identifying a mammalian tumor, useful for predicting a response to HER2-
PT directed therapy, comprises assaying a sample to detect a pattern of

PT expression, phosphorylation or both, of one or more polypeptides.
XX Claim 35; SEQ ID NO 3; 49pp; English.
XX This invention relates to a novel method for determining or predicting
CC the response of a patient to HER-2 directed therapy. Specifically, it
CC refers to analysing a mammalian tumour in order to detect a pattern of
CC expression and/or phosphorylation of a protein taken from the group
CC including insulin growth factor receptor (IGFR) polypeptide, epidermal
CC growth factor receptor (EGFR), phosphorylated S6 ribosomal protein, The
CC phosphorylated AKT, phosphorylated NDF or phosphorylated ERK protein. The
CC present invention describes characterising a mammalian tumour's
CC responsiveness to an HER-2 therapy and hence an individual's response to
CC this cancer therapy, using an immunologically specific antibody directed
CC against one of the aforementioned proteins. Furthermore, it provides
CC predictive biomarkers that can be used to assess the efficacy of
CC therapeutic agents targeted to HER-2/neu. This polypeptide sequence is
CC the human phosphorylated ERK protein of the invention.
XX
SQ Sequence 379 AA;

Query Match 98.2%; Score 1731; DB 8; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKQPDFVGPRTQLOYGEGAYGMVSSAY 60
DB 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKQPDFVGPRTQLOYGEGAYGMVSSAY 60
QY 61 DHVKTTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLKLYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHIL----- 258
DB 241 NRPIFGKHVLDQNLHILGILSPSQEDLNCIIMKARNVLSQPSKTVAWAKLPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEBALHPYLEQYYDPTDEPVABEPTFAMELDDLPKRL 316
DB 301 SKALDLDRLMTFNPKNRITVEBALHPYLEQYYDPTDEPVABEPTFAMELDDLPKRL 360
QY 317 KELIFQETARFQGVLEAP 335
DB 361 KELIFQETARFQGVLEAP 379

RESULT 12
ADV42070
ADV42070 standard; protein; 379 AA.
ADV42070;
24-FEB-2005 (first entry)
Human extracellular signal regulated kinase SEQ ID NO:3.
tumor; antibody therapy; human epidermal growth factor receptor 2; HER2;
phosphorylation; extracellular signal regulated kinase; ERK.
Homo sapiens.
Key Location/Qualifiers
Modified-site 202
Modified-site 204
/note= "phosphorylated"

/note= "phosphorylated"
US2004248151-A1.
09-DEC-2004.
11-DEC-2003; 2003US-00735118.
05-APR-2002; 2002US-0370473P.
11-DEC-2002; 2002US-0432942P.
07-APR-2003; 2003US-00408520.
(VENT-) VENTANA MEDICAL SYSTEMS INC.
(CELL-) CELL SIGNALING TECHNOLOGY.
Bacus SS, Smith BL;
WPI; 2005-072808/08.
REFSEQ; XP_055766.
Identifying mammalian tumor responding to HER2-directed therapy, involves
assaying sample from tumor to detect expression and/or phosphorylation of
insulin-like growth factor receptor polypeptide/erbB1 receptor
polypeptide.
Claim 35; SEQ ID NO 3; 21pp; English.
The invention relates to a novel method for identifying a mammalian tumor
responding to HER2-directed therapy, where the tumor overexpresses HER2.
The method involves assaying a sample obtained from the mammalian tumor
to detect a pattern of expression and/or phosphorylation of one or
several polypeptides chosen from insulin-like growth factor receptor
(IGFR) polypeptide, erbB1 receptor (EGFR) polypeptide, neu
differentiation factor (NDF) polypeptide, phosphorylated S6 ribosomal
polypeptide, phosphorylated AKT polypeptide and phosphorylated
extracellular signal regulated kinase (ERK) polypeptide, where the
detected pattern of expression, phosphorylation or both expression and
phosphorylation identifies mammalian tumors that respond to HER2-directed
therapy or that do not respond to HER2-directed therapy. The method of
the invention is useful for identifying a mammalian tumor that responds
to a HER2-directed therapy or does not respond to HER2-directed therapy,
and for selecting a subject with cancer for treatment with a molecule
targeting HER2 or subject with cancer to not receive treatment with a
molecule targeting HER2. The method is useful for determining or
predicting response to HER2-directed therapy in an individual. The method
enables prediction of response in cancer subjects to cancer therapy. The
present sequence represents the human ERK protein used in the method of
the invention.

Query Match 98.2%; Score 1731; DB 9; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKQPDFVGPRTQLOYGEGAYGMVSSAY 60
DB 1 MAAAAGGGGGPRRTGEGVGPGEVEMVKQPDFVGPRTQLOYGEGAYGMVSSAY 60
QY 61 DHVKTTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTTRVAIKKISPEHQTYCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLKLYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLKLYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHIL----- 258
DB 241 NRPIFGKHVLDQNLHIL----- 258

Db 241 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINMKARNYLQSLPSKTKVNAWAKLPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEALAHPLYEQYYDPTDFVASEPPTFAMELDDLPKRL 316
Db 301 SKALDLDRLMTFNPKNRITVEALAHPLYEQYYDPTDFVASEPPTFAMELDDLPKRL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379
RESULT 13
ADW12905
ID ADW12905 standard; protein; 379 AA.
AC ADW12905;
XX
DT 07-APR-2005 (first entry)
XX
DE Human SNAPK3V3 variant protein.
XX
KW DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW enzyme.
XX
OS Homo sapiens.
XX
PN US2005013817-A1.
XX
PD 20-JAN-2005.
XX
PF 18-JUL-2003; 2003US-00623108.
XX
PR 18-JUL-2003; 2003US-00623108.
XX
PA (DAIK/) DAI K.
XX
PI Dai K;
XX
PI WPI; 2005-080923/09.
DR N-PSDB; ADW12904.
DR
XX
XX New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing
PT diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene
PT in a mammal.
XX
XX Disclosure; SEQ ID NO 6; 55pp; English.
XX
XX The invention relates to novel isolated polypeptide (I) comprising an
CC amino acid sequence selected from sequences comprising 335 or 359 amino
CC acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
CC polynucleotides are useful for diagnosing diseases, e.g. cancers,
CC associated with the deficiency of the SNAPK3 gene in a mammal. The
CC fragments of the polypeptides and polynucleotides can also be used as
CC primers or probes. This sequence corresponds to the SNAPK3V1 variant
CC protein.
XX
XX Sequence 379 AA;
SQ
Query Match 98.2%; Score 1731; DB 9; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAGGGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGYTQYIGEGAGWVSAY 60
Db 1 MAAAAAGGGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGYTQYIGEGAGWVSAY 60
QY 61 DHRKTRVAIKKISPFHQYTCQRTLRBIQILFRHENVIGIRDILFASLTLEAMRDYI 120
Db 61 DHRKTRVAIKKISPFHQYTCQRTLRBIQILFRHENVIGIRDILFASLTLEAMRDYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHI CYFLYQILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHI CYFLYQILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINMKARNYLQSLPSKTKVNAWAKLPKSD 258
Db 241 NRPIFGKHYLDQNLHILGILGPSQEDLNCIINMKARNYLQSLPSKTKVNAWAKLPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEALAHPLYEQYYDPTDFVASEPPTFAMELDDLPKRL 316
Db 301 SKALDLDRLMTFNPKNRITVEALAHPLYEQYYDPTDFVASEPPTFAMELDDLPKRL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379
RESULT 14
ADZ65036
ID ADZ65036 standard; protein; 379 AA.
XX
AC ADZ65036;
XX
DT 14-JUL-2005 (first entry)
XX
DE Mitogen activated protein kinase, MAPK-3, SEQ ID 2.
XX
KW MAP kinase; Mitogen-activated protein kinase; enzyme; cancer.
XX
OS Homo sapiens.
XX
PN US2005095657-A1.
XX
PD 05-MAY-2005.
XX
PF 21-SEP-2004; 2004US-00945684.
XX
PR 11-OCT-2002; 2002US-0418038P.
XX
PR 14-OCT-2003; 2003WO-US032248.
XX
PA (ARBI/) ARBISER J L.
PA (COHE/) COHEN C.
XX
PI Arbisser JL, Cohen C;
XX
XX WPI; 2005-354474/36.
DR REFSEQ; XP_055766.3.
XX
XX Detecting phosphorylated mitogen activated protein kinase (P-MAPK), by
PT contacting sample with antibody having affinity for P-MAPK, detecting
PT antibody/P-MAPK complex to detect P-MAPK, and expression of P-MAPK
PT indicates cancer.
XX
PS Claim 3; SEQ ID NO 2; 23pp; English.
XX
XX The present invention relates to a method (M1) for detecting
CC Phosphorylated Mitogen Activated Protein Kinase (P-MAPK; ADZ65035-
CC ADZ65042), by providing sample, contacting the sample with at least one
CC antibody having affinity for the phosphorylated portion of P-MAPK, and
CC detecting antibody/P-MAPK complex, in which the presence of antibody/P-
CC MAPK complex indicates that P-MAPK is present in sample, and where
CC constitutive expression of P-MAPK is indicative of cancer.
XX
SQ Sequence 379 AA;
Query Match 98.2%; Score 1731; DB 9; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.3e-170;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAGGGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGYTQYIGEGAGWVSAY 60
Db 1 MAAAAAGGGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGYTQYIGEGAGWVSAY 60

QY 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHLYDLQNLHIL----- 258
Db 241 NRPIFGKHLYDLQNLHIL----- 258
QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTTFAMELDDLPKERL 316
Db 301 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTTFAMELDDLPKERL 360
QY 317 KELIFOETARFQPGVLEAP 335
Db 361 KELIFOETARFQPGVLEAP 379
RESULT 15
AAW40540
ID AAW40540 standard; protein; 403 AA.
XX
AC AAW40540;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5471.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's, Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
XX
XX WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
(HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AAI59696.
XX
Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
PT
XX

PS
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAW38642-AAW42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 403 AA;
Query Match 98.2%; Score 1731; DB 4; Length 403;
Best Local Similarity 88.4%; Pred. No. 1.4e-170; Mismatches 0; Indels 44; Gaps 1;
Matches 335; Conservative 0;
QY 1 MAAAAAQQGGGGEPRRTGEGVGFGEVEMVKQPFQVGPRTYQIQYIGEGAYGMVSSAY 60
Db 25 MAAAAAQQGGGGEPRRTGEGVGFGEVEMVKQPFQVGPRTYQIQYIGEGAYGMVSSAY 84
QY 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 85 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 144
QY 121 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 145 VQDLMETDLYKLLKSQQLSNDHICFYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 204
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 205 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 264
QY 241 NRPIFGKHLYDLQNLHIL----- 258
Db 265 NRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLSQSPSKTKVAWAKLPKSD 324
QY 259 --ALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTTFAMELDDLPKERL 316
Db 325 SKALDLDLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVAEPTTFAMELDDLPKERL 384
QY 317 KELIFOETARFQPGVLEAP 335
Db 385 KELIFOETARFQPGVLEAP 403
Search completed: February 6, 2006, 15:29:01
Job time : 274.438 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:29:28 ; Search time 25.1009 Seconds
(without alignments)
1284.124 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAAGGGGGPRRTGV.....LKELIPQETARFQGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*

1: pirl:*

2: pirl:*

3: pirl:*

4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1725	97.8	379	1 A48082	MAP kinase 3 (EC 2
2	1656	93.9	380	2 JC1451	Ca2+/calmodulin-de
3	1655	93.9	377	2 S28184	Ca2+/calmodulin-de
4	1636	92.8	369	2 A60041	Ca2+/calmodulin-de
5	1455.5	82.6	392	2 JW0052	extracellular sign
6	1445.5	82.0	360	2 S25011	protein kinase ERK
7	1444.5	81.9	360	1 JQ1400	MAP kinase 1 (EC 2
8	1434.5	81.4	358	2 S16444	mitogen-activated
9	1434.5	81.4	358	2 A40033	protein kinase (EC
10	1433	81.3	369	2 JW0053	extracellular sign
11	1425	80.8	361	2 A39754	mitogen-activated
12	1293.5	73.4	376	2 A46036	extracellular sign
13	1272	72.1	376	2 A36978	MAP kinase mpk-1 (
14	1272	72.1	444	2 A36977	MAP kinase sur-1 (
15	907	51.4	356	2 T51944	pathogenicity MAP
16	883.5	50.1	361	2 T51943	mitogen-activated
17	855.5	48.5	362	2 A47211	protein kinase ERK
18	844	47.9	415	2 A56042	mitogen-activated
19	840	47.6	372	2 S15663	protein kinase (EC
20	833.5	47.3	387	2 S48123	mitogen-activated
21	828	47.0	394	2 S33635	mitogen-activated
22	821	46.6	393	2 S51321	mitogen-activated
23	820	46.5	395	2 S40472	mitogen-activated
24	808.5	45.9	371	2 S51320	mitogen-activated
25	804	45.6	393	2 T47803	mitogen-activated
26	803.5	45.6	370	2 T47504	mitogen-activated
27	803.5	45.6	370	2 S40469	mitogen-activated
28	802	45.5	376	2 S40470	mitogen-activated
29	801.5	45.5	371	2 T09622	protein kinase MMK

RESULT 1

A48082

MAP kinase 3 (EC 2.7.1.-) - human

N;Alternate names: extracellular signal-regulated kinase 1 (ERK1); mitogen-activated protein kinase (EC 2.7.1.37)

C;Species: Homo sapiens (man)

C;Date: 04-Sep-1998 #sequence revision 04-Sep-1998 #text_change 09-Jul-2004

C;Accession: A48082; PQ0270; S23428; S15519; S21579

R;Charest, D.L.; Mordret, G.; Harder, K.W.; Jirik, F.; Pelech, S.L.

Mol. Cell. Biol. 13, 4679-4690, 1993

A;Title: Molecular cloning, expression, and characterization of the human mitogen-activated protein kinase ERK1

A;Reference number: A48082; MUID:93330262; PMID:7687743

A;Accession: A48082

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-379 <CHA>

A;Cross-references: UNIPROT:P27361; UNIPARC:UPI0000035BE2; EMBL:X60188; NID:g31220; PIDN:AAA36142.1; PID:93330262

A;Experimental source: hepatoma cell line HEP G2

R;Note: authors translated the codon AGC for residue 174 as Ile

R;Owaki, H.; Makar, R.; Boulton, T.G.; Cobb, M.H.; Geppert, T.D.

Biochem. Biophys. Res. Commun. 182, 1416-1422, 1992

A;Title: Extracellular signal-regulated kinases in T cells: characterization of human ERK1

A;Reference number: JQ1400; MUID:92171961; PMID:1540184

A;Accession: PQ0270

A;Molecule type: mRNA

A;Residues: 14-173, 'I', 175-379 <OWA>

A;Cross-references: UNIPARC:UPI000016AB99; GB:M84490; NID:g186695; PIDN:AAA36142.1; PID:93330262

R;Gonzalez, F.A.; Raden, D.L.; Rigby, M.R.; Davis, R.J.

FEBS Lett. 304, 170-178, 1992

A;Title: Heterogeneous expression of four MAP kinase isoforms in human tissues.

A;Reference number: S23426; MUID:92316223; PMID:1319925

A;Accession: S23428

A;Molecule type: mRNA

A;Residues: 25-173, 'I', 175-379 <CON>

A;Cross-references: UNIPARC:UPI000016A2B4; EMBL:Z11696; NID:g23882; PIDN:CAA77754.1; PID:93330262

C;Comment: This enzyme is activated by MAP kinase kinase (see PIR:45100 and PIR:A46723).

C;Genetics:

A;Gene: GDB:PRK3; ERK1

A;Cross-references: GDB:135679; OMIM:601795

A;Map position: 16pter-16qter

C;Complex: monomer

C;Function:

A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine

A;Pathway: MAP kinase cascade

C;Superfamily: kinase-related

C;Keywords: ATP; monomer; phosphoprotein; phosphotransferase; serine/threonine-specific kinase

F;40-330/Domain: protein kinase homology <KIN>

F;48-56/Region: protein kinase ATP-binding motif

F;202/Binding site: phosphate (Thr) (covalent) (by MAP kinase kinase)

F;204/Binding site: phosphate (Tyr) (covalent) (by MAP kinase kinase)

ALIGNMENTS

Query Match		97.8%;	Score 1725;	DB 1;	Length 379;
Best Local Similarity		88.1%;	Pred. No. 1.8e-70;		
Matches 334;		Conservative 0;	Mismatches 1;	Indels 44;	Gaps 1;
QY	1	MAAAAQGGGGGPRTEGVGPGEVEMVKGQPPDVGPVRYTQLOYIGEGAYGMVSSAY	60		
Db	1	MAAAAQGGGGGPRTEGVGPGEVEMVKGQPPDVGPVRYTQLOYIGEGAYGMVSSAY	60		
QY	61	DHVKTRVAIKKISPEHOTYQORTLREIQILLRPHENVIGIRDILRASTLEAMRDVIV	120		
Db	61	DHVKTRVAIKKISPEHOTYQORTLREIQILLRPHENVIGIRDILRASTLEAMRDVIV	120		
QY	121	VDLMDTLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVLRDLKPSNLLINTTCDL	180		
Db	121	VDLMDTLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVLRDLKPSNLLINTTCDL	180		
QY	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
Db	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
QY	241	NRPIFGKHVLDLQNLHIL-----	258		
Db	241	NRPIFGKHVLDLQNLHILGILGSPQEDLNCIINMKARNYLSLPKTKVAVAKLPPKSD	300		
QY	259	--ALDLLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPERLK	316		
Db	301	SKALDLLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPERLK	360		
QY	317	KELIFOETARFQGVLEAP 335			
Db	361	KELIFOETARFQGVLEAP 379			
RESULT 2					
JC1451					
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) - rat					
N/Alternate names: ERK1-MAP kinase; extracellular signal-regulated kinase					
C/Species: Rattus norvegicus (Norway rat)					
C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004					
C/Accession: JC1451; A35061; A37140; A40466; S24947					
R/Marquardt, B.; Stabel, S.					
Gene 120, 297-299, 1992					
A/Title: Sequence of a rat cDNA encoding the ERK1-MAP kinase.					
A/Reference number: JC1451; MUID:93013050; PMID:1327976					
A/Accession: JC1451					
A/Molecule type: mRNA					
A/Residues: 1-380 <WAR>					
A/Cross-references: UNIPROT:P21708; UNIPARC:UPI000012F174; EMBL:X65198; NID:g56626; PIDN					
A/Experimental source: brain					
R/Boulton, T.G.; Yancopoulos, G.D.; Gregory, J.S.; Slaughter, C.; Moomaw, C.; Hsu, J.; C					
Science 249, 64-67, 1990					
A/Title: An insulin-stimulated protein kinase similar to yeast kinases involved in cell					
A/Reference number: A35061; MUID:90312137; PMID:2164259					
A/Accession: A35061					
A/Molecule type: mRNA					
A/Residues: 14-380 <BOU>					
A/Cross-references: UNIPARC:UPI0000145072; GB:M38194; NID:g204051; PIDN:AAA41123.1; PID:					
R/Boulton, T.G.; Gregory, J.S.; Cobb, M.H.					
Biochemistry 30, 278-286, 1991					
A/Title: Purification and properties of extracellular signal-regulated kinase 1, an insu					
A/Reference number: A37140; MUID:91105092; PMID:1846291					
A/Accession: A37140					
A/Status: preliminary					
A/Molecule type: protein					
A/Residues: 43-64;167-178, 'X',180-183, 'X',185 <BO2>					
A/Cross-references: UNIPARC:UPI000017558F; UNIPARC:UPI0000175590					
R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.					
DNA Cell Biol. 10, 505-514, 1991					
A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse a					
A/Reference number: A40466; MUID:91369479; PMID:1716439					
A/Accession: A40466					
A/Status: preliminary					
A/Molecule type: mRNA					

A/Residues: 14-94, 'R',96-380 <DEM>

A/Cross-references: UNIPARC:UPI00001707CE; GB:U12008; GB:S59509; NID:g515498; PIDN:AAA20C

C/Superfamily: kinase-related transforming protein; protein kinase homology

C/Keywords: ATP; calmodulin binding; phosphoprotein; phosphotransferase; serine/threonine

F:41-331/Domain: protein kinase homology <Kin>

F:49-57/Region: protein kinase ATP-binding motif

Query Match	93.9%;	Score 1656;	DB 2;	Length 380;
Best Local Similarity	85.4%;	Pred. NO. 2.1e-67;		
Matches 323;	Conservative 1;	Mismatches 10;	Indels 44;	Gaps 1;

QY	2	AAAAAQQGGGGPRTEGVGPGEVEMVKGQPPDVGPVRYTQLOYIGEGAYGMVSSAYD	61		
Db	3	AAAAAAPGGGGGPRGTAGVVPVPGVEVAVKGQPPDVGPVRYTQLOYIGEGAYGMVSSAYD	62		
QY	62	HVKTRVAIKKISPEHOTYQORTLREIQILLRPHENVIGIRDILRASTLEAMRDVIV	121		
Db	63	HVKTRVAIKKISPEHOTYQORTLREIQILLRPHENVIGIRDILRASTLEAMRDVIV	122		
QY	122	QDLMDTLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVLRDLKPSNLLINTTCDLK	181		
Db	123	QDLMDTLYKLLKSQOLSNNDHICYFLYQILRLKYIHSANVLRDLKPSNLLINTTCDLK	182		
QY	182	ICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	241		
Db	183	ICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	242		
QY	242	RFIFGKHVLDLQNLHIL-----	258		
Db	243	RFIFGKHVLDLQNLHILGILGSPQEDLNCIINMKARNYLSLPKTKVAVAKLPPKSDS	302		
QY	259	-ALDLLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPERLK	317		
Db	303	KALDLLDRMLTFNPNKRITVBEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPERLK	362		
QY	318	ELIFQETARFQGVLEAP	335		
Db	363	ELIFQETARFQGAPEAP	380		

RESULT 3

S28184

Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) erk-1 - mouse (fragment)

N/Alternate names: extracellular signal-regulated kinase 1

C/Species: Mus musculus (house mouse)

C/Date: 22-Nov-1993 #sequence revision 30-Jan-1998 #text_change 09-Jul-2004

C/Accession: S28184; B40466; A41371

R/Tanner, B.; Mueckler, M.

Biochim. Biophys. Acta 1171, 319-320, 1993

A/Title: Molecular cloning of a mouse extracellular signal regulated kinase (erk-1).

A/Reference number: S28184; MUID:93144347; PMID:8424957

A/Accession: S28184

A/Molecule type: mRNA

A/Residues: 1-377 <TAN>

A/Cross-references: UNIPROT:Q63844; UNIPARC:UPI000017558C; EMBL:Z14249

R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.

DNA Cell Biol. 10, 505-514, 1991

A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse

A/Reference number: A40466; MUID:91369479; PMID:1716439

A/Accession: B40466

A/Molecule type: mRNA

A/Residues: 74-377 <DEM>

A/Cross-references: UNIPARC:UPI0000023599; GB:S59517

R/Crews, C.M.; Alessandrini, A.A.; Erikson, R.L.

Proc. Natl. Acad. Sci. U.S.A. 88, 8845-8849, 1991

A/Title: Mouse Erk-1 gene product is a serine/threonine protein kinase that has the potential to

A/Reference number: A41371; MUID:92020947; PMID:1717989

A/Accession: A41371

A/Molecule type: mRNA

A/Residues: 7-16 <CRE>

A/Cross-references: UNIPARC:UPI0000170C57; GB:S58470; NID:g236372; PIDN:AAB19973.1; PID:AA20C

C/Superfamily: kinase-related transforming protein; protein kinase homology

C/Keywords: ATP; calmodulin binding; phosphotransferase; protein kinase

S25011

Db

protein kinase ERK2 (EC 2.7.1.-) - bovine
N:Alternate names: extracellular signal-regulated kinase
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S25011
R:Ely, C.M.; Cox, M.E.; Her, J.; Parsons, S.J.
submitted to the EMBL Data Library, July 1992
A:Description: Cloning and sequencing of ERK2 from a bovine adrenal medulla cDNA library
A:Reference number: S25011
A:Accession: S25011
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-360 <ELY>
A:Cross-references: UNIPROT:P46196; UNIPARC:UPI000012F172; EMBL:Z14089; NID:g3336; PIDN:C
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:23-313/Domain: protein kinase homology <KIN>
F:31-39/Region: protein kinase ATP-binding motif
Query Match 82.0%; Score 1445.5; DB 2; Length 360;
Best Local Similarity 74.6%; Pred. No. 4.6e-58;
Matches 279; Conservative 17; Mismatches 17; Indels 61; Gaps 2;
Db 1 MAAAAAAGAG-----PEMVRGQVDFVGPVTEGVPGEVVMKQDPDVGPRYTQIQYIGEGAYGMVSSAY 60
QY 1 MAAAAAAGAG-----PEMVRGQVDFVGPVTEGVPGEVVMKQDPDVGPRYTQIQYIGEGAYGMVSSAY 60
Db 1 MAAAAAAGAG-----PEMVRGQVDFVGPVTEGVPGEVVMKQDPDVGPRYTQIQYIGEGAYGMVSSAY 60
QY 61 DHVTRKTRVAIKKISPEHQTQRTLEIQTILRRPHENVIGIRDIRASTLEAMRDVYI 120
Db 44 DNVKVRVAIKKISPEHQTQRTLEIKILRRPHENIIGINDIIRAPTEQMKDVI 103
QY 121 VDLMETDLYKLLKSOQLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 180
Db 104 VDLMETDLYKLLKTOHLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 163
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 164 KICDFGLARVADPDHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 223
QY 241 NRPIPGKHYLDQNLHIL-----PEMVRGQVDFVGPVTEGVPGEVVMKQDPDVGPRYTQIQYIGEGAYGMVSSAY 258
Db 224 NRPIPGKHYLDQNLHILGILGPSQEDLNCIINKARNYLLSLPHKNKVPWNLFPNAD 283
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 316
Db 284 SKALDLDLKMFTFNPKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMELDDLPKEKL 343
QY 317 KELIFOETARFQPG 330
Db 344 KELIFEETARFQPG 357
RESULT 7
JQ1400
MAP kinase 1 (EC 2.7.1.-) - human
N:Alternate names: ERK; extracellular signal-regulated kinase 2 (ERK2); mitogen-activate
N:Contains: protein kinase (EC 2.7.1.37)
C:Species: Homo sapiens (man)
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
C:Accession: JQ1400; S23426; S23427; S21577; S21578
R:Owaki, H.; Makar, R.; Boulton, T.G.; Cobb, M.H.; Gerecht, T.D.
Biochem. Biophys. Res. Commun. 192, 1416-1422, 1992
A:Title: Extracellular signal-regulated kinases in T cells: characterization of human ER
A:Reference number: JQ1400; MUID:92171961; PMID:1540184
A:Accession: JQ1400
A:Molecule type: mRNA
A:Residues: 1-360 <OWA>
A:Cross-references: UNIPROT:P28482; UNIPARC:UPI000000104P; GB:M84489; NID:g182190; PIDN:
A:Experimental source: cell line Jurkat
R:Gonzalez, F.A.; Raden, D.L.; Rigby, M.R.; Davis, R.J.
FEBS Lett. 304, 170-178, 1992
A:Title: Heterogeneous expression of four MAP kinase isoforms in human tissues.
A:Reference number: S23426; MUID:92316223; PMID:1319925

A:Accession: S23426
A:Molecule type: mRNA
A:Residues: 13-360 <GON1>
A:Cross-references: UNIPARC:UPI000016A28D; EMBL:Z11695; NID:g23878; PIDN:CAA77753.1; PID:
A:Accession: S23427
A:Molecule type: mRNA
A:Residues: 1-90, 'O', '92-360 <GON2>
A:Cross-references: UNIPARC:UPI000016A299; EMBL:Z11694; NID:g23880; PIDN:CAA77752.1; PID:
C:Comment: This enzyme is activated by MAP kinase kinase (see PIR:45100 and PIR:A46723).
C:Genetics:
A:Gene: GDB:PRKM1; ERK; MAPK1; ERK2
A:Cross-references: GDB:135677; OMIM:176948
A:Map position: 22q11.2-22q11.2
C:Complex: monomer
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
A:Pathway: MAP kinase cascade
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F:23-313/Domain: protein kinase homology <KIN>
F:31-39/Region: protein kinase ATP-binding motif
F:185/Binding site: phosphate (Thr) (covalent) (by MAP kinase kinase) #status predicted
F:187/Binding site: phosphate (Tyr) (covalent) (by MAP kinase kinase) #status predicted
Query Match 81.9%; Score 1444.5; DB 1; Length 360;
Best Local Similarity 74.3%; Pred. No. 5.1e-58;
Matches 278; Conservative 18; Mismatches 17; Indels 61; Gaps 2;
QY 1 MAAAAAAGAG-----PEMVRGQVDFVGPVTEGVPGEVVMKQDPDVGPRYTQIQYIGEGAYGMVSSAY 60
Db 1 MAAAAAAGAG-----PEMVRGQVDFVGPVTEGVPGEVVMKQDPDVGPRYTQIQYIGEGAYGMVSSAY 60
QY 61 DHVTRKTRVAIKKISPEHQTQRTLEIQTILRRPHENVIGIRDIRASTLEAMRDVYI 120
Db 44 DNVKVRVAIKKISPEHQTQRTLEIKILRRPHENIIGINDIIRAPTEQMKDVI 103
QY 121 VDLMETDLYKLLKSOQLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 180
Db 104 VDLMETDLYKLLKTOHLSNDHICVFLYQILRGLKYIHSANVLRDLKPSNLLINTCDL 163
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 164 KICDFGLARVADPDHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 223
QY 241 NRPIPGKHYLDQNLHIL-----PEMVRGQVDFVGPVTEGVPGEVVMKQDPDVGPRYTQIQYIGEGAYGMVSSAY 258
Db 224 NRPIPGKHYLDQNLHILGILGPSQEDLNCIINKARNYLLSLPHKNKVPWNLFPNAD 283
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKERL 316
Db 284 SKALDLDLKMFTFNPKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMELDDLPKEKL 343
QY 317 KELIFOETARFQPG 330
Db 344 KELIFEETARFQPG 357
RESULT 8
S16444
N:Alternate names: protein kinase (EC 2.7.1.-) - mouse
N:Alternate names: cell division cycle-2-related protein kinase ERK2
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
C:Accession: S16444; S15015; FN0480
R:Her, J.H.; Wu, J.; Rall, T.B.; Sturgill, T.W.; Weber, M.J.
Nucleic Acids Res. 19, 3743, 1991
A:Title: Sequence of pp42/MAP kinase, a serine/threonine kinase regulated by tyrosine ph
A:Reference number: S16444; MUID:91305126; PMID:1649458
A:Accession: S16444
A:Molecule type: mRNA
A:Residues: 1-358 <HER>
A:Cross-references: UNIPROT:P27703; UNIPARC:UPI0000003F8A; EMBL:X58712; NID:g53001; PIDN:
A:Note: part of this sequence was confirmed by protein sequencing

R;Payne, D.M.; Rossomando, A.J.; Martino, P.; Erickson, A.K.; Her, J.H.; Shabanowitz, J.
EMBO J. 10, 885-892, 1991
A;Title: Identification of the regulatory phosphorylation sites in pp42/mitogen-activated
A;Reference number: S15015; MUID:91184134; PMID:1849075
A;Accession: S15015
A;Status: preliminary
A;Molecule type: protein
A;Residues: 171-181, 'X', 183-189 <P>AY>
A;Cross-references: UNIPARC:UPI0000175588
R;Erehler, M.A.; Nagorekaya, T.V.; Visser, J.W.M.; Belyavsky, A.V.
Gene 124, 305-306, 1993
A;Title: Novel CDC2-related protein kinases produced in murine hematopoietic stem cells.
A;Reference number: PN0479; MUID:93185941; PMID:8444355
A;Accession: PN0480
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 145-195 <ERS>
A;Cross-references: UNIPARC:UPI0000175589
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine
F;21-311/Domain: protein kinase homology <KIN>
F;29-37/Region: protein kinase ATP-binding motif
F;183/Binding site: phosphate (Thr) (covalent) #status experimental
F;185/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 81.4%; Score 1434.5; DB 2; Length 358;
Best Local Similarity 73.8%; Pred. No. 1.4e-57;
Matches 276; Conservative 19; Mismatches 16; Indels 63; Gaps 2;

QY 1 MAAAAAGGGGGEPRRTGEGVPGVGEVEMVKGQPPFDVGPRTYQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAG-----PEMVRGQVFDVGPRTYNSLYIGEGAYGMVCSAY 41
QY 61 DHVKTTRVAIKKISPFPHQTYCQRTLEIQLIRFRHENVIGIRDLRASTLEAMRDVYI 120
DB 42 DNLNKRVAIKKISPFPHQTYCQRTLEIKILLFRHENIIGINDIIRAPTIQOMKDVYI 101
QY 121 VQDLMETDLYKLLKSQOLSNDDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 102 VQDLMETDLYKLLKTOHLSNDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 161
QY 181 KICDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 162 KICDFGLARVADPHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 221
QY 241 NRPIFGKHYLDQNLHIL----- 258
DB 222 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINLKARNYLLSLPHKNKVPWNLFPNAD 281
QY 259 --ALDLDRLMTFNPKNRTVEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKRL 316
DB 282 SKALDLDLKMFTNPKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMELDDLPKRL 341
QY 317 KELIFOETARFQPG 330
DB 342 KELIFEETARFQPG 355

RESULT 10
JW0053
extracellular signal-regulated kinase (EC 2.7.-.-) 2 - common carp
N;Alternate names: CERK2
C;Species: Cyprinus carpio (common carp)
C;Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
C;Accession: JW0053
R;Hashimoto, H.; Yokoyama, Y.; Matsuo, Y.; Toyohara, H.; Kohno, M.; Sakaguchi, M.
J. Biochem. 123, 1031-1035, 1998
A;Title: Existence of two isoforms of extracellular signal-regulated kinase in fish.
A;Reference number: JW0052; MUID:98269030; PMID:9603989
A;Accession: JW0053
A;Molecule type: mRNA
A;Residues: 1-369 <HAS>
A;Cross-references: UNIPROT:Q7LZH5; UNIPARC:UPI0000175597
C;Comment: This enzyme is the key enzyme in the intracellular signal transduction pathway
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase
F;32-322/Domain: protein kinase homology <KIN>

Query Match 81.3%; Score 1433; DB 2; Length 369;
Best Local Similarity 74.3%; Pred. No. 1.7e-57;
Matches 277; Conservative 17; Mismatches 25; Indels 54; Gaps 3;

QY 2 AAAAAAGGGGGEPRRTGEGVPGVGEVEMVKGQPPFDVGPRTYQLOVIGEGAYGMVSSAYD 61
DB 4 AAVSAPAGGGPNP-----GSG-----AEMVRGQAFDVGPRYTNLSYIGEGAYGMVCSAYK 53
QY 62 HVAKTTRVAIKKISPFPHQTYCQRTLEIQLIRFRHENVIGIRDLRASTLEAMRDVYIV 121
DB 54 RDNKRVVAIKKISPFPHQTYCQRTLEIKILLFRHENIIGINDIIRTTTIDQMKDVYIV 113
QY 122 QDLMETDLYKLLKSQOLSNDDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK 181
DB 114 QDLMETDLYKLLKTOHLSNDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK 173

R;Payne, D.M.; Rossomando, A.J.; Martino, P.; Erickson, A.K.; Her, J.H.; Shabanowitz, J.
EMBO J. 10, 885-892, 1991
A;Title: Identification of the regulatory phosphorylation sites in pp42/mitogen-activated
A;Reference number: S15015; MUID:91184134; PMID:1849075
A;Accession: S15015
A;Status: preliminary
A;Molecule type: protein
A;Residues: 171-181, 'X', 183-189 <P>AY>
A;Cross-references: UNIPARC:UPI0000175588
R;Erehler, M.A.; Nagorekaya, T.V.; Visser, J.W.M.; Belyavsky, A.V.
Gene 124, 305-306, 1993
A;Title: Novel CDC2-related protein kinases produced in murine hematopoietic stem cells.
A;Reference number: PN0479; MUID:93185941; PMID:8444355
A;Accession: PN0480
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 145-195 <ERS>
A;Cross-references: UNIPARC:UPI0000175589
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine
F;21-311/Domain: protein kinase homology <KIN>
F;29-37/Region: protein kinase ATP-binding motif
F;183/Binding site: phosphate (Thr) (covalent) #status experimental
F;185/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 81.4%; Score 1434.5; DB 2; Length 358;
Best Local Similarity 73.8%; Pred. No. 1.4e-57;
Matches 276; Conservative 19; Mismatches 16; Indels 63; Gaps 2;

QY 1 MAAAAAGGGGGEPRRTGEGVPGVGEVEMVKGQPPFDVGPRTYQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAG-----PEMVRGQVFDVGPRTYNSLYIGEGAYGMVCSAY 41
QY 61 DHVKTTRVAIKKISPFPHQTYCQRTLEIQLIRFRHENVIGIRDLRASTLEAMRDVYI 120
DB 42 DNLNKRVAIKKISPFPHQTYCQRTLEIKILLFRHENIIGINDIIRAPTIQOMKDVYI 101
QY 121 VQDLMETDLYKLLKSQOLSNDDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 102 VQDLMETDLYKLLKTOHLSNDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 161
QY 181 KICDFGLARIADPHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 162 KICDFGLARVADPHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 221
QY 241 NRPIFGKHYLDQNLHIL----- 258
DB 222 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINLKARNYLLSLPHKNKVPWNLFPNAD 281
QY 259 --ALDLDRLMTFNPKNRTVEALAHPLYEQYDPTDEPVAEPTTFAMELDDLPKRL 316
DB 282 SKALDLDLKMFTNPKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMELDDLPKRL 341
QY 317 KELIFOETARFQPG 330
DB 342 KELIFEETARFQPG 355

RESULT 9
A40033
protein kinase (EC 2.7.1.37) ERK2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004
C;Accession: A40033
R;Boulton, T.G.; Nye, S.H.; Robbins, D.J.; Ip, N.Y.; Radziejewska, E.; Morgenbesser, S.D.
Cell 65, 663-675, 1991
A;Title: ERKs: a family of protein-serine/threonine kinases that are activated and tyros
A;Reference number: A40033; MUID:91235302; PMID:2032290
A;Accession: A40033
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 <BOU>
A;Cross-references: UNIPROT:P27703; UNIPARC:UPI0000003FEA; GB:M64300; NID:g204055; PIDN:

```
QY 182 ICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 241
DB 174 ICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 233
QY 242 RPIFGKHYLDQNLNHL----- 258
DB 234 RPIFGKHYLDQNLNHLILGILGPSQEDLNCIINIKARNYLLSLPLCKYVNNRLEFPNADP 293
QY 259 -ALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERLJK 317
DB 294 KALDLLDKMLTFNPHKRIVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKETLJK 353
QY 318 ELIFQETARFQPG 330
DB 354 ELIFEETARFQPG 366

RESULT 11
A39754
mitogen-activated protein kinase (EC 2.7.1.1-) - African clawed frog
N:Alternate names: M phase MAP kinase
C:Species: Xenopus laevis (African clawed frog)
C>Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 09-Jul-2004
C:Accession: A39754; S16597
R:Posada, J.; Sanghera, J.; Pelech, S.; Aebersold, R.; Cooper, J.A.
Mol. Cell. Biol. 11, 2517-2528, 1991
A:Title: Tyrosine phosphorylation and activation of homologous protein kinases during oocyte maturation in Xenopus oocytes
A:Reference number: A39754; MUID:91203872; PMID:1708093
A:Accession: A39754
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-361 <POS>
A:Cross-references: UNIPROT:P26696; UNIPARC:UPI0000175592; GB:M60977
R:Gotoh, Y.; Moriyama, K.; Matsuda, S.; Okumura, E.; Kishimoto, T.; Kawasaki, H.; Suzuki, EMO J. 10, 2661-2668, 1991
A:Title: Xenopus M phase MAP kinase: isolation of its cDNA and activation by MPF.
A:Reference number: S16597; MUID:91330892; PMID:1714387
A:Accession: S16597
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4,'G','6','A','8-28','1','30-31','A','33-46','D','48','V','50-331','F',333-361 <GOT>
A:Cross-references: UNIPARC:UPI000017156A; EMBL:X59813; NID:G64893; PIDN:CAA42482.1; PIDN:CAA42482.1; PIDN:CAA42482.1
C:Superfamily: kinase-related transforming protein: protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:26-316/Domain: protein kinase homology <KIN>
F:34-42/Region: protein kinase ATP-binding motif

Query Match 80.8%; Score 1425; DB 2; Length 361;
Best Local Similarity 73.5%; Pred. No. 3.7e-57;
Matches 275; Conservative 18; Mismatches 23; Indels 58; Gaps 2;

QY 1 MAAAAAGGGGGPRRTGPGVGPGEVEMVKGQPFVGPVGTQLOQYIGEGAYGMVSSAY 60
DB 1 MAAAAASSNPGGSP-----EMVRQAQFDVGPVGTNLSYIGEGAYGMVCSAH 46
QY 61 DHVKTQVIAKKISPEHQYTCQRTREIQILLRPHENVIGIRDLRASTLEAMRDVYI 120
DB 47 CNINKVRVAIKKISPEHQYTCQRTREIKILLRPHENIIGINDIRAPTEQMKDVYI 106
QY 121 VQDLMTDLYKLLKSQOLSNLHCYFLYQILRLKVIHNSANVLRDLKPSNLLINTTCDL 180
DB 107 VQDLMTDLYKLLKTOHLSNDHICYFLYQILRLKVIHNSANVLRDLKPSNLLINTTCDL 166
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 240
DB 167 KICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS 226
QY 241 NRPIFGKHYLDQNLNHL----- 258
DB 227 NRPIFGKHYLDQNLNHLILGILGPSQEDLNCIINIKARNYLLSLPLCKYVNNRLEFPNAD 286
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 316
```

```
DB 287 PKALDLLDKMLTFNPHKRIVEEALAHPLYEQYDPSPEVAEAPLKFEMLDDLPKETL 346
QY 317 KELIFQETARFQPG 330
DB 347 KELIFEETARFQPG 360

RESULT 12
A46036
extracellular signal-regulated kinase (EC 2.7.1.1-) DmERK-A - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46036
R:Biggs III, W.H.; Zipursky, S.L.
Proc. Natl. Acad. Sci. U.S.A. 89, 6295-6299, 1992
A:Title: Primary structure, expression, and signal-dependent tyrosine phosphorylation of Drosophila melanogaster ERK-A
A:Reference number: A46036; MUID:92335284; PMID:1378625
A:Accession: A46036
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-376 <BIG>
A:Cross-references: UNIPROT:P40417; UNIPARC:UPI00000000FB3; GB:M95124; NID:G485754; PIDN:G485754
A>Note: sequence extracted from NCBI backbone (NCBI:108684, NCBI:108685)
C:Genetics:
A:Gene: FlyBase:rl
A:Cross-references: FlyBase:FBgn0003256
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase
F:36-326/Domain: protein kinase homology <KIN>
F:44-52/Region: protein kinase ATP-binding motif

Query Match 73.4%; Score 1293.5; DB 2; Length 376;
Best Local Similarity 71.7%; Pred. No. 2.7e-51;
Matches 251; Conservative 22; Mismatches 32; Indels 45; Gaps 2;

QY 24 VP-GEVEMVKGQPFVGPVGTQLOQYIGEGAYGMVSSAYDHVKTQVIAKKISPEHQYTC 82
DB 19 VPQSNAEVIRGQIFEVGPVYIKLAIIGEGAYGMVVSADDTLLNQVIAKKISPEHQYTC 78
QY 83 QSTLEIRIQILLRPHENVIGIRDLRASTLEAMRDVYIVQDLMTDLYKLLKSQOLSNL 142
DB 79 QSTLEIRITLTFKHENIIDIRDLRASTLEAMRDVYIVQDLMTDLYKLLKSQOLSNL 138
QY 143 ICYFLYQILRLKVIHNSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFLT 202
DB 139 ICYFLYQILRLKVIHNSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFLT 198
QY 203 EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFGKHYLDQNLNHI----- 257
DB 199 EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFGKHYLDQNLNHI 258
QY 258 -----LALDLLDRMLTFNPNKRITVE 278
DB 259 SPSRDDLECIINEKARNYLLSLPLCKYVNNRLEFPNADALDLGKMLTFNPHKRIPVE 318
QY 279 EALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERLKEIFQETARFQ 328
DB 319 EALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERLKEIFQETARFQ 368

RESULT 13
A36978
MAP kinase mpk-1 (EC 2.7.1.1-) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A36978; T22118
R:Lackner, M.R.; Kornfeld, K.; Miller, L.M.; Horvitz, H.R.; Kim, S.K.
Genes Dev. 8, 160-173, 1994
A:Title: A MAP kinase homolog, mpk-1, is involved in ras-mediated induction of vulval cell division
A:Reference number: A36978; MUID:94131270; PMID:8299936
A:Accession: A36978
A>Status: preliminary
```

A;Molecule type: mRNA
A;Residues: 1-376 <LAC>
A;Cross-references: UNIPROT:P39745; UNIPARC:UPI000002B465; GB:U271124; GB:S68854; NID:986
R;Jassal, B.; Smith, A.
submitted to the EMBL Data Library, December 1994
A;Reference number: Z19519
A;Accession: T22118
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-376 <WTL>
A;Cross-references: UNIPARC:UPI000002B465; EMBL:Z46937; PIDN:CAA87057.1; GSPDB:GN000021;
A;Experimental source: clone F43C1
C;Genetics:
A;Gene: mpk-1
A;Map position: 3
A;Introns: 23/3; 43/2; 85/3; 104/2; 304/2
A;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;26-285/Domain: protein kinase homology <KIN>
F;34-42/Region: protein kinase ATP-binding motif

Query Match 72.1%; Score 1272; DB 2; Length 376;
Best Local Similarity 70.0%; Pred. No. 2.5e-50;
Matches 238; Conservative 33; Mismatches 25; Indels 44; Gaps 1;

Qy 28 VEMVKGQFPDVGPRYTQLOYIGEGAYGVSSAYDHYVTRKTRVAIKKISPFPHQTYCQRTL 87
Db 14 VEEVHQLFEVAPRYVNLVIGEGAYGVASALDTITRDRVAIKKISPFPHQTYCQRTL 73

Qy 88 EIQILLRPHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKLSQQLSNDHCYFL 147
Db 74 EIKILNRFKHENIINIOEIRSETSDIKDIYVQCLMETDLYKLKLTQKLSNDHCYFL 133

Qy 148 YQILRGUKYTHSANVLRDLKPSNLLNTTCDLKICDFGLARIADPEHDHTGFLTEYVAT 207
Db 134 YQILRGUKYTHSANVLRDLKPSNLLNTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 193

Qy 208 RWTYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRPIFFPGKHLYDQLNHL----- 258
Db 194 RWTYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRPLFPFGKHLYDQLNHLAVVGSNSNA 253

Qy 259 -----ALDILDRMLTFNPNKRITVEEALAH 283
Db 254 DLQCIINDKARSYLIISLPHKPKQPWARYPGADPRALDLDKMLTFNPNHRIDIEQALAH 313

Qy 284 PYLEQYVDPTDEPVAREPFTFAMELDDLPKRLKELIFOE 323
Db 314 PYLEQYVDGDEPVCCEPFTLEMEFDDLPKRLKELIWE 353

RESULT 14
A36977
MAP kinase sur-1 (EC 2.7.1.-) - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: A36977
R;Wu, Y.; Han, M.
Genes Dev. 8, 147-159, 1994
A;Title: Suppression of activated Let-60 ras protein defines a role of Caenorhabditis el
A;Reference number: A36977; MUID:94131269; PMID:8299935
A;Accession: A36977
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-444 <WUA>
A;Cross-references: UNIPROT:P39745; UNIPARC:UPI0000164181; GB:U03879; NID:G433356; PIDN:
C;Genetics:
A;Gene: sur-1
A;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;94-353/Domain: protein kinase homology <KIN>
F;102-110/Region: protein kinase ATP-binding motif

Query Match 72.1%; Score 1272; DB 2; Length 444;

Best Local Similarity 70.0%; Pred. No. 2.9e-50;
Matches 238; Conservative 33; Mismatches 25; Indels 44; Gaps 1;

Qy 28 VEMVKGQFPDVGPRYTQLOYIGEGAYGVSSAYDHYVTRKTRVAIKKISPFPHQTYCQRTL 87
Db 82 VEEVHQLFEVAPRYVNLVIGEGAYGVASALDTITRDRVAIKKISPFPHQTYCQRTL 141

Qy 88 EIQILLRPHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKLSQQLSNDHCYFL 147
Db 142 EIKILNRFKHENIINIOEIRSETSDIKDIYVQCLMETDLYKLKLTQKLSNDHCYFL 201

Qy 148 YQILRGUKYTHSANVLRDLKPSNLLNTTCDLKICDFGLARIADPEHDHTGFLTEYVAT 207
Db 202 YQILRGUKYTHSANVLRDLKPSNLLNTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 261

Qy 208 RWTYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRPIFFPGKHLYDQLNHL----- 258
Db 262 RWTYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRPLFPFGKHLYDQLNHLAVVGSNSNA 321

Qy 259 -----ALDILDRMLTFNPNKRITVEEALAH 283
Db 322 DLQCIINDKARSYLIISLPHKPKQPWARYPGADPRALDLDKMLTFNPNHRIDIEQALAH 381

Qy 284 PYLEQYVDPTDEPVAREPFTFAMELDDLPKRLKELIFOE 323
Db 382 PYLEQYVDGDEPVCCEPFTLEMEFDDLPKRLKELIWE 421

RESULT 15
T51944
pathogenicity MAP kinase 1 [imported] - Pyricularia grisea
C;Species: Pyricularia grisea
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T51944
R;Xu, J.R.; Hamer, J.E.
Genes Dev. 10, 2696-2706, 1996
A;Title: MAP kinase and CAMP signaling regulate infection structure formation and pathog
A;Reference number: Z25880; MUID:97102671; PMID:8946911
A;Accession: T51944
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-356 <XU>
A;Cross-references: UNIPROT:Q92246; UNIPARC:UPI000006C824; EMBL:U70134; PIDN:AAC49521.2
A;Experimental source: strain Guy11
C;Genetics:
A;Gene: PMK1
A;Introns: 39/2; 232/1
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 51.4%; Score 907; DB 2; Length 356;
Best Local Similarity 52.2%; Pred. No. 4.1e-34;
Matches 177; Conservative 47; Mismatches 67; Indels 48; Gaps 2;

Qy 36 FDVGPRTYQLOYIGEGAYGVSSAYDHYVTRKTRVAIKKISPFPHQTYCQRTLRIQILRF 95
Db 18 FNVSEQYDQDVVGEAGYGVCSAIHKPSGQKVAIKKITPFDHSMFCLRLIREKLLRYF 77

Qy 96 RHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKLSQQLSNDHCYFLYQILRLGLK 155
Db 78 NHENIISILDIOKPRSEVFNEVYLQELMETDHRVIRTDLSDDDHCQFYIOTLRALK 137

Qy 156 YTHSANVLRDLKPSNLLNTTCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAPEI 215
Db 138 AMHSANVLRDLKPSNLLNANCDLVCFGLARSAAQENNSGFMTEYVATRWYRAPEI 197

Qy 216 MLNSGYTKSIDWSVGCILAEMLSNRPIFFPGKHLYDQLNHL----- 257
Db 198 MLTFKSYTKAIDVMSVGCILAEMLSKGPLFPFGKHLYDQLNHL----- 257

Qy 258 -----LALDILDRMLTFNPNKRITVEEALAHPLRYQYD 291
Db 258 RAREYIRSLPFFKKVPRTLPFKTSDLAUDLLEKLAFAFPVKITVEEALKHPYLPYHD 317

QY 292 PTDEPVA---ERPFTFAMELDDLPKXRLKELIFOETAR 326
Db 318 PDDEPTAPPPEEFFDFDKHKNLSKEQLKQFIYQEIIR 356

Search completed: February 6, 2006, 15:35:33
Job time : 26.1009 secs

Qy	1	MAAAAAQGGGGGPPRRTEGVGPVGEVEMVKGQPFVDGPRYTQLOIYIGEGAYGMVSSAY	60
Db	1	MAAAAAQGGGGGPPRRTEGVGPVGEVEMVKGQPFVDGPRYTQLOIYIGEGAYGMVSSAY	60
Qy	61	DHVKTVAIKKISPFESHQTYCQRTLRRIQILLFRHENVIGIRDILRASTLEAMRDVVI	120
Db	61	DHVKTVAIKKISPFESHQTYCQRTLRRIQILLFRHENVIGIRDILRASTLEAMRDVVI	120
Qy	121	VQDLMETDLYKLAKKSQQLSNDHICVFYLIQILRGUKYIHSANVLHRDLKPSNLLINTTCDL	180

Db 121 VQDLMETDLYKLLKSQSLNDHICYFLYQILRGLKVHISANVLHRLDKPSNLLSNTTCDL 180
 Qy 181 KICDFGLARIADPHDHTGFLTYVATRWYRAPEIMLNSKGVTYSIDWSVGCILAEML 240
 Db 181 KICDFGLARIADPHDHTGFLTYVATRWYRAPEIMLNSKGVTYSIDWSVGCILAEML 240
 Qy 241 NRPIFGKHVLDQNLHILADLLDMLTFNPNKRITVEALAHPLYLEQYVDPTDEPVAEE 300
 Db 241 NRPIFGKHVLDQNLHILADLLDMLTFNPNKRITVEALAHPLYLEQYVDPTDEPVAEE 300
 Qy 301 PPTFAMELDDLPKRLKELIFQSTARFQGVLEAP 335
 Db 301 PPTFAMELDDLPKRLKELIFQSTARFQGVLEAP 335

RESULT 2
 MK03_HUMAN
 ID MK03_HUMAN STANDARD; PRT; 379 AA.
 AC Pw7361;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
 DE signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
 DE (MAP kinase 1) (MAPK 1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
 DE associated protein 2 kinase)
 DE Names=MAPK3; Synonyms=ERK1, PRK3;
 GN Homo sapiens (Human).
 OS
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC NUCLEOTIDE SEQUENCE.
 RP TISSUE=Hepatoma;
 RX MEDLINE=93330262; PubMed=7687743;
 RA Charest D.L., Jirik P., Harder K., Pelech S.L., Mordret G.;
 RT "Molecular cloning, expression, and characterization of the human
 RT mitogen-activated protein kinase p44erk1.";
 RL Mol. Cell. Biol. 13:4679-4690(1993).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uudin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 14-379.
 RX MEDLINE=92171961; PubMed=1540184;
 RA Owaki H., Makar R., Boulton T.G., Cobb M.H., Geppert T.D.;
 RT "Extracellular signal-regulated kinases in T cells: characterization
 RT of human ERK1 and ERK2 cDNAs.";
 RL Biochem. Biophys. Res. Commun. 182:1416-1422(1992).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 25-379.

RX MEDLINE=92316223; PubMed=1319925; DOI=10.1016/0014-5793(92)80612-K;
 RA Gonzalez F.A., Raden D.L., Rigby M.R., Davis R.J.;
 RT "Heterogeneous expression of four MAP kinase isoforms in human
 RT tissues.";
 RL FEBS Lett. 304:170-178(1992).
 RN [5]
 RP INTERACTION WITH HIV-1 NEF.
 RX PubMed=8794306;
 RA Greenway A.L., Azad A., Mills J., McPhee D.A.;
 RT "Human immunodeficiency virus type 1 Nef binds directly to LCK and
 RT mitogen-activated protein kinase, inhibiting kinase activity.";
 RL J. Virol. 70:6701-6708(1996).
 CC !- FUNCTION: Involved in both the initiation and regulation of
 CC meiosis, mitosis, and postmitotic functions in differentiated
 CC cells by phosphorylating a number of transcription factors such as
 CC Elk-1. Phosphorylates E1F4EBP1; required for initiation of
 CC translation. Phosphorylates microtubule-associated protein 2
 CC (MAP2) (By similarity).
 CC !- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC !- COFACTOR: Magnesium (By similarity).
 CC !- ENZYME REGULATION: Activated and tyrosine phosphorylated in
 CC response to insulin and NGF. This interaction inhibits its kinase
 CC !- SUBUNIT: Binds to HIV-1 Nef. This interaction inhibits its kinase
 CC activity.
 CC !- INTERACTION:
 CC P53355:DAPK1; NbExp=3; IntAct=EBI-73995, EBI-358616;
 CC O75676:RP96K44; NbExp=1; IntAct=EBI-73995, EBI-73933;
 CC !- PTM: Autophosphorylated on threonine and tyrosine residues (By
 CC similarity).
 CC !- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
 CC kinase subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; X60188; CAA42744.1; -; mRNA.
 DR EMBL; BC013992; AAH13992.1; -; mRNA.
 DR EMBL; M84490; AAA36142.1; -; mRNA.
 DR EMBL; Z11696; CAA77754.1; -; mRNA.
 DR PIR; A48082; A48082.
 DR HSP; P28482; 1PME.
 DR SMR; P27361; 29-374.
 DR IntAct; P27361; -.
 DR Ensembl; ENSG00000102882; Homo sapiens.
 DR HGNC; HGNC:6877; MAPK3.
 DR H-InvDB; HIX0012930; -.
 DR Reactome; P27361; -.
 DR MIM; 601795; -.
 DR GO; GO:0005524; F-ATP binding; NAS.
 DR GO; GO:0004707; F-MAP kinase activity; NAS.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 DR GO; GO:0000074; P:regulation of cell cycle; NAS.
 DR InterPro; IPR008349; Erk 1-2 MAPK.
 DR InterPro; IPR008351; JNK_MAPK.
 DR InterPro; IPR003527; MAP_kin.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser Thr_pkin_AS.
 DR InterPro; IPR002290; Ser Thr_pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR01770; ERK1ERK2MAPK.
 DR PRINTS; PR01772; JNKMAPKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; cell cycle; Kinase; Nucleotide-binding; Phosphorylation;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 42 330 Protein kinase.


```
FT NP_BIND 48 56 ATP (By similarity).
FT ACT_SITE 166 166 Proton acceptor (By similarity).
FT BINDING 71 71 ATP (By similarity).
FT MOD_RES 202 202 Phosphothreonine (activates the kinase).
FT MOD_RES 204 204 Phosphotyrosine (activates the kinase).
FT CONFLICT 174 174 I -> S (in Ref. 1).
SQ SEQUENCE 379 AA; 43136 MW; B6020CE413EC41F7 CRC64;

Query Match 98.2%; Score 1731; DB 1; Length 379;
Best Local Similarity 88.4%; Pred. No. 1-7e-100;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRTEGVGPGVEVMVKQDPDVGPRYTQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRTEGVGPGVEVMVKQDPDVGPRYTQLOVIGEGAYGMVSSAY 60

QY 61 DHVYKTRVAIKKISPEHQYTCORTLEIQILRPHENVIGIRDIRLRASTLEAMRDVYI 120
Db 61 DHVYKTRVAIKKISPEHQYTCORTLEIQILRPHENVIGIRDIRLRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLKKSQOLSNDHICYFLYQILRGLKITHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLKKSQOLSNDHICYFLYQILRGLKITHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTYVATRWTRAPFIMLSNGYTKSIDIVSGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTYVATRWTRAPFIMLSNGYTKSIDIVSGCILAEMLS 240

QY 241 NRPIFGKHVLDQNLHIL----- 258
Db 241 NRPIFGKHVLDQNLHIL----- 258

QY 259 --ALDLDLRLMTFNPKRIITVEALAHPLYEQYDPTDEPVAEPEPTFAMELDDLPKRL 316
Db 301 SKALDLDLRLMTFNPKRIITVEALAHPLYEQYDPTDEPVAEPEPTFAMELDDLPKRL 360

QY 317 KELIFQETARFQGVLEAP 335
Db 361 KELIFQETARFQGVLEAP 379

RESULT 3
MK03 MOUSE STANDARD; PRT; 380 AA.
ID _MK03 MOUSE STANDARD; PRT; 380 AA.
AC Q63844; O61531; ORK0X5; O91W5;
DT 01-NOV-1997 (Rel. 35, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
DE (MAP kinase 1) (MAPK1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
DE associated protein 2 kinase) (MNK1).
GN Name=Mapk3; Synonyms=Erk1, Pmk3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=FVB/N; TISSUE=Kidney, and Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Stachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Dapkin M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Bosak S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

```
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 4-380.
RX PubMed=8424957; DOI=10.1016/0167-4781(93)90074-N;
RA Tanner B., Mueckler M.;
RT "Molecular cloning of a mouse extracellular signal regulated kinase
RT (erk-1).";
RL Biochim. Biophys. Acta 1171:319-320 (1993).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 10-19.
RC TISSUE=Pre-B cell;
RX MEDLINE=92020947; PubMed=1717989;
RA Crews C.M., Alessandrini A.A., Erikson R.L.;
RT "Mouse Erk-1 gene product is a serine/threonine protein kinase that
RT has the potential to phosphorylate tyrosine.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8845-8849 (1991).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 77-380.
RC TISSUE=Fetal brain;
RX MEDLINE=91369479; PubMed=1716439;
RA de Miguel C., Kligman D., Patel J., Detex-Wadleigh S.D.;
RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA
RT from mouse and rat brain.";
RL DNA Cell Biol. 10:505-514 (1991).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 171-209.
RC STRAIN=CBA; TISSUE=Bone marrow;
RX MEDLINE=93185941; PubMed=8444355; DOI=10.1016/0378-1119(93)90411-U;
RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
RT "Novel CDC2-related protein kinases produced in murine hematopoietic
RT stem cells.";
RL Gene 124:305-306 (1993).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 171-205.
RX MEDLINE=93092802; PubMed=1459009;
RA Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;
RT "Identification of new protein kinase genes, similar to kinases of the
RT cdc2 family and expressed in murine hematopoietic stem cells.";
RL Dokl. Akad. Nauk SSSR 324:893-897 (1992).
CC -1- FUNCTION: Involved in both the initiation and regulation of
CC meiosis, mitosis, and postmitotic functions in differentiated
CC cells by phosphorylating a number of transcription factors such as
CC ELK-1. Phosphorylates ERF4EBP1; required for initiation of
CC translation. Phosphorylates microtubule-associated protein 2
CC (MAP2) (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
CC phosphorylation.
CC -1- PTM: Autophosphorylated on threonine and tyrosine residues.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC kinase subfamily.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC ENBL; BC013754; AAH13754.1; -; mRNA.
CC ENBL; BC029712; AAH29712.1; -; mRNA.
CC ENBL; S58470; AAB19973.1; -; mRNA.
CC ENBL; X64605; CAA45889.1; -; mRNA.
CC PIR; S28184; S28184.
CC HSSP; P28482; IPME.
```

SMR; Q63844; 30-375.
 DR Ensembl; ENSMUSG00000063065; Mus musculus.
 DR MGI; MGI:1346859; Mapk3.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0004704; F:MAP kinase activity; IMP.
 DR GO; GO:0001784; F:phosphotyrosine binding; IMP.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0004672; F:protein kinase activity; IDA.
 DR GO; GO:0009887; F:organogenesis; IDA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 DR GO; GO:0006974; P:response to DNA damage stimulus; IDA.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR008349; Erk 1-2 MAPK.
 DR InterPro; IPR008351; JNK_MAPK.
 DR InterPro; IPR003527; MAP_kin.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR0008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR01770; ERK1ERK2MAPK.
 DR PRINTS; PR01772; JNKMAPKINASE.
 DR ProDom; PD000001; JNKMAPKINASE.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Cell cycle; Kinase; Nucleotide-binding; Phosphorylation;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 43 331 Protein kinase.
 FT NP_BIND 49 57 ATP (By similarity).
 FT ACT_SITE 167 167 Proton acceptor (By similarity).
 FT BINDING 72 72 ATP (By similarity).
 FT MOD_RES 203 203 Phosphothreonine (activates the kinase)
 FT MOD_RES 205 205 Phosphotyrosine (activates the kinase)
 FT CONFLICT 178 178 T -> P (in Ref. 5 and 6).
 FT SEQUENCE 380 AA; 43066 MW; 49C14A95B627237F CRC64;
 Query Match 94.1%; Score 1659; DB 1; Length 380;
 Best Local Similarity 85.4%; Pred. No. 5.4e-96;
 Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;
 QY 2 AAAAAAGGGGPRRTGCGVPGVGVEMVKGQPPDVGPRTYQLQVIGEGAGVSSAYD 61
 DB 3 AAAAAAGGGGPRGTAGVVPVGVGVVKGQPPDVGPRTYQLQVIGEGAGVSSAYD 62
 QY 62 HVKTRVAIKTSPRHQTYCQRTLEIOILLFRHENVIGIRDILRASTLEAMRDVYIV 121
 DB 63 HVKTRVAIKTSPRHQTYCQRTLEIOILLFRHENVIGIRDILRASTLEAMRDVYIV 122
 QY 122 QDLMETDLYKLKSKQSLNDHCYFLYQLIRGLKYIHSANVLHRLDKPSNLLINTTCDLK 181
 DB 123 QDLMETDLYKLKSKQSLNDHCYFLYQLIRGLKYIHSANVLHRLDKPSNLLINTTCDLK 182
 QY 182 ICDFGLARIADPHDTGFLTEVATRWTRAPBIMLSNGYKTSIDINWSGCIILAEMLSN 241
 DB 183 ICDFGLARIADPHDTGFLTEVATRWTRAPBIMLSNGYKTSIDINWSGCIILAEMLSN 242
 QY 242 RPIFFGCKHYLDQLNHIL----- 258
 DB 243 RPIFFGCKHYLDQLNHILGILGPSQSDLNLCINMKARNYLQSLPSTKVAVAKLFPKSDS 302
 QY 259 -ALDLDRLMTFNPKNKRTIVVEALAHPLYEQYDDTDEPVAEPPFTFAMELDDLPKERLK 317
 DB 303 KALDLDRLMTFNPKNKRTIVVEALAHPLYEQYDDTDEPVAEPPFTFAMELDDLPKERLK 362
 QY 318 ELIFQETARFQGVLEAP 335
 DB 363 ELIFQETARFQGAPEGP 380

RESULT 4
 MK03 RAT
 ID MK03 RAT STANDARD; PRT; 380 AA.
 AC F21708; Q62686; Q9JUL3;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
 signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
 DE (MAP kinase 1) (MAPK 1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
 associated protein 2 kinase) (MNK1).
 GN Name=Mapk3; Synonyms=Erk1, Prkm3;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1] NUCLEOTIDE SEQUENCE.
 RP STRAIN=Sprague-Dawley; TISSUE=Brain;
 RC MEDLINE=93013050; PubMed=1327976; DOI=10.1016/0378-1119(92)90109-3;
 RA Marquardt B., Stabel S.;
 RT "Sequence of a rat cDNA encoding the ERK1-MAP kinase.";
 RL Gene 120:297-299(1992).
 RN [2] NUCLEOTIDE SEQUENCE OF 7-380.
 RP MAISONPIERRE P.C., le Beau M.M., Espinosa R. III, Ip N.Y.,
 RA Belluscio L., la Monte S.M., Squinto S., Furth M.E., Yancopoulos G.D.;
 RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
 RN [3] NUCLEOTIDE SEQUENCE OF 14-380, AND PARTIAL PROTEIN SEQUENCE.
 RX MEDLINE=90312137; PubMed=2164259;
 RA Boulton T.G., Yancopoulos G.D., Gregory J.S., Slaughter C., Moomaw C.,
 RA Hsu J., Cobb M.H.;
 RT "An insulin-stimulated protein kinase similar to yeast kinases
 involved in cell cycle control.";
 RL Science 249:64-67(1990).
 RN [4] NUCLEOTIDE SEQUENCE OF 14-380.
 RP TISSUE=Brain cortex;
 RX MEDLINE=91369479; PubMed=1716439;
 RA de Miguel C., Kilgman D., Patel J., Detera-Wadleigh S.D.;
 RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA
 from mouse and rat brain.";
 RL DNA Cell Biol. 10:505-514(1991).
 RN [5] NUCLEOTIDE SEQUENCE (ISOFORM B).
 RX MEDLINE=20283605; PubMed=10748187; DOI=10.1074/jbc.M910060199;
 RA Yung Y., Yao Z., Hancock T., Seger R.;
 RT "ERK1b, a 46-kDa ERK isoform that is differentially regulated by
 MEK.";
 RL J. Biol. Chem. 275:15799-15808(2000).
 RN [6] PROTEIN SEQUENCE OF 43-64 AND 167-185, AND CHARACTERIZATION.
 RX MEDLINE=91105092; PubMed=1846291;
 RA Boulton T.G., Gregory J.S., Cobb M.H.;
 RT "Purification and properties of extracellular signal-regulated kinase
 1, an insulin-stimulated microtubule-associated protein 2 kinase.";
 RL Biochemistry 30:278-286(1991).
 RN [7] AUTOPHOSPHORYLATION.
 RX MEDLINE=91296777; PubMed=1712480;
 RA Seger R., Ahn N.G., Boulton T.G., Yancopoulos G.D., Panayotatos N.,
 RA Radziejewski E., Ericsson L., Bratlien R.L., Cobb M.H., Krebs E.G.;
 RT "Microtubule-associated protein 2 kinases, ERK1 and ERK2, undergo
 autophosphorylation on both tyrosine and threonine residues:
 implications for their mechanism of activation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:6142-6146(1991).
 RN [8] PHOSPHORYLATION OF EIF4EBP1.
 RP MEDLINE=95025978; PubMed=7939721;
 RA Lin T.-A., Kong X., Haystead T.A.J., Pause A., Belsham G.J.,

RA Sonenberg N., Lawrence J.C. Jr.;
 RT "PHAS-I as a link between mitogen-activated protein kinase and
 translation initiation."; Science 266:653-656(1994).
 RL
 CC -1- FUNCTION: Involved in both the initiation and regulation of
 CC meiosis, mitosis, and postmitotic functions in differentiated
 CC cells by phosphorylating a number of transcription factors such as
 CC Elk-1. Phosphorylates E1F4BP1; required for initiation of
 CC translation. Phosphorylates microtubule-associated protein 2
 CC (MAP2) (By similarity).
 CC
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC
 CC -1- COFACTOR: Magnesium (By similarity).
 CC -1- ENZYME REGULATION: Activated and tyrosine phosphorylated in
 CC response to insulin and NGF.
 CC -1- SUBCELLULAR LOCATION: Nuclear (isoform B).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=A;
 CC IsoId=P21708-1; Sequence=Displayed;
 CC Name=B;
 CC IsoId=P21708-2; Sequence=VSP_004830;
 CC TISSUE SPECIFICITY: Highest levels within the nervous system,
 CC expressed in different tissues, mostly in intestine, placenta and
 CC lung.
 CC -1- DEVELOPMENTAL STAGE: Increased expression during development.
 CC -1- PTM: Autophosphorylated on threonine and tyrosine residues.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
 CC kinase subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 DR EMBL; X65198; CAA46318.1; -; mRNA.
 DR EMBL; M61177; AAA63486.1; -; mRNA.
 DR EMBL; M38194; AAA41123.1; -; mRNA.
 DR EMBL; U12008; AAA20009.1; -; mRNA.
 DR EMBL; AF155236; AAF71666.1; -; mRNA.
 DR PIR; JCI451; JCI451.
 DR HSSP; P28482; IPME.
 DR SNR; P21708; 30-375.
 DR Ensembl; ENSRNOG0000019601; Rattus norvegicus.
 DR RGD; 3046; Mapk3.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR GO; GO:0005826; C:insoluble fraction; IDA.
 DR GO; GO:0004707; F:MAP kinase activity; TAS.
 DR GO; GO:0000189; P:nuclear translocation of MAPK; IDA.
 DR GO; GO:0045941; P:positive regulation of transcription; IEF.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 DR GO; GO:0007243; P:protein kinase cascade; TAS.
 DR InterPro; IPR008345; Erk_1_2_MAPK.
 DR InterPro; IPR008351; JNK_MAPK.
 DR InterPro; IPR003527; MAP kin.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR008271; Ser Thr pkin AS.
 DR InterPro; IPR002290; Ser Thr pkinase.
 DR Pfam; PF00069; Pkinase; I.
 DR PRINTS; PR01770; ERK1ERK2MAPK.
 DR PRINTS; PR01772; JNKMAPKINASE.
 DR ProDom; PD000001; Prot kinase; I.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW Alternative splicing; ATP-binding; Cell cycle;
 KW Direct protein sequencing; Kinase; Nuclear protein;
 KW Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase;
 KW Transferase.
 FT DOMAIN 43 331 Protein kinase.
 FT NP_BIND 49 57 ATP (By similarity).

FT	ACT_SITE	167	167	Proton acceptor (By similarity).
FT	BINDING	72	72	ATP (By similarity).
FT	MOD_RES	203	203	Phosphothreonine (activates the kinase)
FT				(By similarity).
FT	MOD_RES	205	205	Phosphotyrosine (activates the kinase)
FT				(By similarity).
FT	VARSPLIC	340	340	E -> EVSRPPAAGRGISVPSVPRVPYCICPQ (in
FT				isoform B).
FT	CONFLICT	95	95	/FTId=VSP_004830.
FT				G -> R (in Ref. 4).
SQ	SEQUENCE	380 AA;	42981 MW;	53785C4ED722723A CRC64;

Query Match 93.9%; Score 1656; DB 1; Length 380;
 Best Local Similarity 85.4%; Pred. No. 8.3e-96;
 Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;

QY	2	AAAAAGGGGGPRRTGEGVGPVGEVMVKGQPFDPVGPRTYQLOQYIGGAYGMVSSAYD	61
DB	3	AAAAAPGGGGPRGTAGVVPVGEVEVKGQPFDPVGPRTYQLOQYIGGAYGMVSSAYD	62
QY	62	HVKTRVAIKISPEEHQTYCQRTLRRIQILRFHENVIGIRDIRASTLEAMRDYIV	121
DB	63	HVKTRVAIKISPEEHQTYCQRTLRRIQILRFHENVIGIRDIRASTLEAMRDYIV	122
QY	122	QDLMETDYLKLSQQLSNDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK	181
DB	123	QDLMETDYLKLSQQLSNDHCYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK	182
QY	182	ICDFGLARTADPEHDTGFLTYVATRVAPRPEIMLSNGYTKSIDWSVGCILAEMLSN	241
DB	183	ICDFGLARTADPEHDTGFLTYVATRVAPRPEIMLSNGYTKSIDWSVGCILAEMLSN	242
QY	242	RIPIFGKHYLDQLNHIL	258
DB	243	RIPIFGKHYLDQLNHIL	302
QY	259	-ALDLDRLMTFNPKNRITVEEALAHPLYEYDPTDEPVAEPPFFAMELDDLPERLK	317
DB	303	KALDLDRLMTFNPKNRITVEEALAHPLYEYDPTDEPVAEPPFFAMELDDLPERLK	362
QY	318	ELIFQETARFQPGVLEAP 335	
DB	363	ELIFQETARFQPGVLEAP 380	

RESULT 5
 Q4PIY8 RAT
 ID Q4PIY8_RAT PRELIMINARY; PRT; 380 AA.
 AC Q4PIY8;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Extracellular-signal-regulated kinase 1.
 GN Name=Mapk3;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE-93013050; PubMed-1327976; DOI=10.1016/0378-1119(92)90109-3;
 RA Marquardt B., Stabel S.;
 RT "Sequence of a rat cDNA encoding the ERK1-MAP kinase.";
 RL Gene 120:297-299(1992).
 DR ENBL; S46779; AAA11604.1; -; mRNA.
 KW Kinase.
 SQ SEQUENCE 380 AA; 42981 MW; 53785C4ED722723A CRC64;

Query Match 93.9%; Score 1656; DB 2; Length 380;
 Best Local Similarity 85.4%; Pred. No. 8.3e-96;
 Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;

QY 2 AAAAAAGGGGGERTEGVPVGEVEMVKQPFVDPVRYTQLOVIGEGAYGMVSSAYD 61
Db 3 AAAAAAGGGGGERCTAGVVPVGEVEMVKQPFVDPVRYTQLOVIGEGAYGMVSSAYD 62
QY 62 HVKTRVAIKKISPFHQYTCQRTLREIQILRRFHHENVIGIRDLIRASTLEAMRDVYIV 121
Db 63 HVKTRVAIKKISPFHQYTCQRTLREIQILRRFHHENVIGIRDLIRASTLEAMRDVYIV 122
QY 122 QDLMETDLYKLLSQQLSNDHICVFLYQILRGKYIHSANVLHRDLKPSNLLINTTCDLK 181
Db 123 QDLMETDLYKLLSQQLSNDHICVFLYQILRGKYIHSANVLHRDLKPSNLLINTTCDLK 182
QY 182 ICDFGLARTADPHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN 241
Db 183 ICDFGLARTADPHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN 242
QY 242 RPIFFPGKHVLDQNLHIL----- 258
Db 243 RPIFFPGKHVLDQNLHILGSPSQEDLNCIINMKARNVLSLPSKTKVAAKLPKSDS 302
QY 259 -ALDLDRLMTFNPKNRITVEEALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLK 317
Db 303 KALDLDRLMTFNPKNRITVEEALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLK 362
QY 318 ELIFQETARFQGVLEAP 335
Db 363 ELIFQETARFQGAPEAP 380

RESULT 6
Q7M0H9 CRIGR
ID Q7M0H9 CRIGR PRELIMINARY; PRT; 369 AA.
AC Q7M0H9
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) ERK1 (Fragment).
DE Cricetulus griseus (Chinese hamster).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Cricetidae; Cricetinae; Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92199340; PubMed=1372523;
RA Meloche S., Pages G., Pouyssegur J.;
RT "Functional expression and growth factor activation of an epitope-tagged p44 mitogen-activated protein kinase, p44(map).";
RL Mol. Biol. Cell 3:63-71(1992).
CC -1- FUNCTION: Kinase involved in a signal transduction pathway that is activated by changes in the osmolarity of the extracellular environment. Plays an essential role in maintaining water homeostasis (By similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC PIR; A60041; A60041.
DR HSSP; Q16539; IOZI.
DR SMR; Q7M0H9; 19-364.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k. . .; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR00719; Prof_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; I.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase; 1.

DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-Binding;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
FT NON_TER 369 369
SQ SEQUENCE 369 AA; 42126 MW; 0BB8FCAAC690B4FF CRC64;

Query Match 92.8%; Score 1636; DB 2; Length 369;
Best Local Similarity 85.6%; Pred. No. 1.4e-94;
Matches 316; Conservative 2; Mismatches 7; Indels 44; Gaps 1;

QY 11 GSEPRTEGVGPGVGEVEMVKQPFVDPVRYTQLOVIGEGAYGMVSSAYDHVRKTRVAI 70
Db 1 GGEPRGAAGVGPVGEVEMVKQPFVDPVRYTQLOVIGEGAYGMVSSAYDHVRKTRVAI 60
QY 71 KKISPFHQYTCQRTLREIQILRRFHHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLY 130
Db 61 KKISPFHQYTCQRTLREIQILRRFHHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLY 120
QY 131 KLLKSQQLSNDHICVFLYQILRGKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARI 190
Db 121 KLLKSQQLSNDHICVFLYQILRGKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARI 180
QY 191 ADPEHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSNRPFPKGHY 250
Db 181 ADPEHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSNRPFPKGHY 240
QY 251 LDQLNLHIL-----ALDLDRLM 266
Db 241 LDQLNLHILGSPSQEDLNCIINMKARNVLSLPSKTKVAAKLPKSDSKALDLDRLM 300
QY 267 LTFNPNKRITVEEALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLKLIQETAR 326
Db 301 LTFNPNKRITVRDALAHPLYEQYDTPDEPVAEPPFTFAMELDDLKPKERLKLIQETAR 360
QY 327 FQGVLEAP 335
Db 361 FQGAPEAP 369

RESULT 7
Q8NHX1 HUMAN
ID Q8NHX1 HUMAN PRELIMINARY; PRT; 357 AA.
AC Q8NHX1
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-related kinase 1b.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Aebersold D.M., Yung Y., Seger R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033607; AAK52329.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q8NHX1; 29-344.
DR Ensemble; ENSG00000102882; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR01770; ERK1ERK2MAPK.
 DR PRINTS; PR01772; JNKMAPKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS01017; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00111; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Nucleotide-binding;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 357 AA; 40062 MW; 58C277398ADA79 CRC64;
 Query Match 86.4%; Score 1522.5; DB 2; Length 357;
 Best Local Similarity 85.8%; Pred. No. 1.8e-87;
 Matches 296; Conservative 1; Mismatches 3; Indels 45; Gaps 2;
 QY 1 MAAAAAGGGGGGPRRTGEGVGPGEVEMVKGQPFVDPVPRYTQYIGEGAYGMVSSAY 60
 DB 1 MAAAAAGGGGGGPRRTGEGVGPGEVEMVKGQPFVDPVPRYTQYIGEGAYGMVSSAY 60
 QY 61 DHVRKTRVAIKKISPPFHQTYCQRTLRERIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
 DB 61 DHVRKTRVAIKKISPPFHQTYCQRTLRERIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
 QY 121 VDQIMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 DB 121 VDQIMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
 QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDNSVGCILAEMLS 240
 DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDNSVGCILAEMLS 240
 QY 241 NRPIFGKHLYDQNLNHL----- 258
 DB 241 NRPIFGKHLYDQNLNHLGILGSPSQEDLNCIINMKARNYQLSLPSKTKVAKLPPKSD 300
 QY 259 --ALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPP 301
 DB 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDE-VGQSP 344
 RESULT 8
 ID Q7LZH4_CYPCA PRELIMINARY; PRT; 392 AA.
 AC Q7LZH4_2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Extracellular signal-regulated kinase (EC 2.7.-.-) 1 (Extracellular
 DE signal regulated protein kinase 1).
 GN Name=cERK1;
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Ovary;
 RX MEDLINE=98269030; PubMed=9603989;
 RA Hashimoto H., Yokoyama Y., Matsuo Y., Toyohara H., Kohno M.,
 RA Sakaguchi M.;
 RT "Existence of two isoforms of extracellular signal-regulated kinase in
 RT fish.";
 RL J. Biochem. 123:1031-1035(1998).
 CC -|- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AB006038; BAD23842.1; -; mRNA.
 DR PIR; J00052; J00052.
 DR HSSP; Q16539; 1021.
 DR SMR; Q7LZH4; 39-386.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004707; F:MAP kinase activity; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR008349; Erk_1_2_MAPK.
 DR InterPro; IPR003527; MAP_kin.
 DR InterPro; IPR00719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR01770; ERK1ERK2MAPK; 2.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS01017; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00111; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Nucleotide-binding;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 392 AA; 44133 MW; 02BE90132B5C347B CRC64;
 Query Match 82.6%; Score 1455.5; DB 2; Length 392;
 Best Local Similarity 77.3%; Pred. No. 3.1e-83;
 Matches 286; Conservative 13; Mismatches 19; Indels 52; Gaps 4;
 QY 3 AAAAAAGGGGGGPRRTGEGVGPGEVEMVKGQPFVDPVPRYTQYIGEGAYGMVSSAYDH 62
 DB 25 AAVAVGGANG-----AAGP-KPG-LESVKGNQFVDPVPRYTQYIGEGAYGMVCSAFDN 76
 QY 63 VVKTRVAIKKISPPFHQTYCQRTLRERIQILLRFRHENVIGIRDIRASTLEAMRDVYI 122
 DB 77 VVKTRVAIKKISPPFHQTYCQRTLRERIQILLRFRHENVIGIRDIRASTLEAMRDVYI 136
 QY 123 DQIMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI 182
 DB 137 DQIMETDLYKLLKSQQLSNDHICFYLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI 196
 QY 183 CDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDNSVGCILAEMLSNR 242
 DB 197 CDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDNSVGCILAEMLSNR 256
 QY 243 PIFPGKHLYDQNLNHL----- 258
 DB 257 PIFPGKHLYDQNLNHLGILGSPSQEDLNCIINMKARNYQLSLPSKTKVAKLPPKADNK 316
 QY 259 ALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFAMELDDLPKERLKE 318
 DB 317 ALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFAMELDDLPKERLKE 376
 QY 319 LIQFETARFQ 328
 DB 377 LIYEETARFQ 386
 RESULT 9
 ID Q8UMG6_CHICK PRELIMINARY; PRT; 368 AA.
 AC Q8UMG6_2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Extracellular signal-regulated kinase 2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Pinel gland;
 RA Kasahara T., Hagiashi C., Okano T., Fukada Y.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -|- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AY033635; AAK56503.1; -; mRNA.
 DR HSSP; P28482; 1PME
 DR SMR; Q8UMG6; 24-367.
 DR Ensembl; ENSGALG000000001501; Gallus gallus.


```

QY 3 AAAAOGGGGPRRTTEGVGPGVGVEMVKGPFDVGPRTYTOLOQYIGEGAYGMVSSAYDH 62
Db 18 SSAAGPGGAVAPGGPSGAAGSKFG-LESVKGNQFVGPRTYDLOQYIGEGAYGMVCSAFDN 76
QY 63 VRKTRVAIKKISPFHEQTYCQRTLEIRIQLLRPHENHVGIRDIRLRASTLEAMRDVYIVQ 122
Db 77 VNKIRVAIKKISPFHEQTYCQRTLEIRIKILLRPHENHIIINDILRARHIDYMRDVIYVQ 136
QY 123 DLMETDLYKLLKSQOLSNHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI 182
Db 137 DLMETDLYKLLKTKQOLSNHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI 196
QY 183 CDFGLARIADPEHDHTGFLTEYVATRYAPRAPEIMLSKGYTKSIDWSVGCILAEMLSNR 242
Db 197 CDFGLARIADPEHDHTGFLTEYVATRYAPRAPEIMLSKGYTKSIDWSVGCILAEMLSNR 256
QY 243 PIFPGKHLYDLQNLHIL----- 258
Db 257 PIFPGKHLYDLQNLHILGSPSQDILNCIINMKARNYLQSLPQPKIPWNKLPKADNK 316
QY 259 ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVAEBPFTFAMELDDLPERLKE 318
Db 317 ALDLDRLMTFNPKNRINVEQALAHPLYEQYDPSDEPVAEBPFTFNMELDDLPERLKE 376
QY 319 LIFOETARFQ 328
Db 377 LIFEETARFQ 386

```

RESULT 11

```

Q9DGR6 BRARE
ID Q9DGR6 BRARE PRELIMINARY; PRT; 391 AA.
AC Q9DGR6
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ERK1.
GN Name=mapk3; Synonyms=zERK1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hirano T., Fujii R., Hibi M.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030902; BAB11812.1; -; mRNA.
DR HSP; P28482; 1PME.
DR SMR; Q9DGR6; 40-385.
DR ZFIN; ZDR-GENE-040121-1; mapk3.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk 1.2 MAPK.
DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 391 AA; 44126 MW; 6638F0807B5F9889 CRC64;

```

```

Query Match 82.1%; Score 1447.5; DB 2; Length 391;
Best Local Similarity 76.5%; Pred. No. 9.6e-83;
Matches 283; Conservative 14; Mismatches 28; Indels 45; Gaps 2;

QY 3 AAAAOGGGGPRRTTEGVGPGVGVEMVKGPFDVGPRTYTOLOQYIGEGAYGMVSSAYDH 62
Db 17 SSAAGPGGAVAPGGPSGAAGSKFG-LESVKGNQFVGPRTYDLOQYIGEGAYGMVCSAFDN 75
QY 63 VRKTRVAIKKISPFHEQTYCQRTLEIRIQLLRPHENHVGIRDIRLRASTLEAMRDVYIVQ 122
Db 76 VNKIRVAIKKISPFHEQTYCQRTLEIRIKILLRPHENHIIINDILRARHIDYMRDVIYVQ 135
QY 123 DLMETDLYKLLKSQOLSNHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI 182
Db 136 DLMETDLYKLLKTKQOLSNHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKI 195
QY 183 CDFGLARIADPEHDHTGFLTEYVATRYAPRAPEIMLSKGYTKSIDWSVGCILAEMLSNR 242
Db 196 CDFGLARIADPEHDHTGFLTEYVATRYAPRAPEIMLSKGYTKSIDWSVGCILAEMLSNR 255
QY 243 PIFPGKHLYDLQNLHIL----- 258
Db 256 PIFPGKHLYDLQNLHILGSPSQDILNCIINMKARNYLQSLPQPKIPWNKLPKADNK 315
QY 259 ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVAEBPFTFAMELDDLPERLKE 318
Db 316 ALDLDRLMTFNPKNRINVEQALAHPLYEQYDPSDEPVAEBPFTFNMELDDLPERLKE 375
QY 319 LIFOETARFQ 328
Db 376 LIFEETARFQ 385

```

RESULT 12

```

MK01_BOVIN
ID MK01_BOVIN STANDARD; PRT; 360 AA.
AC P46196;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase
DE 2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERK1).
DE Name=MAPK1; Synonyms=ERK2, PRKM1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ely C.M., Cox M.E., Her J., Parsons S.J.;
RT "Cloning and sequencing of ERK2 from a bovine adrenal medulla cDNA
RT library.";
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Phosphorylates microtubule-associated protein 2 (MAP2).
CC Myelin basic protein (MBP), and Elk-1; may promote entry in the
CC cell cycle.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium (by similarity).
CC -1- ENZYME REGULATION: Activated by phosphorylation on tyrosine and
CC threonine in response to insulin and NGF.
CC -1- PTM: Autophosphorylated on threonine and tyrosine residues, which
CC correlates with a slow and low level of autoactivation (By
CC similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC kinase subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not

```


EMBL; M84489; AA58459.1; -; mRNA.
 DR EMBL; Z11694; CAA7752.1; -; mRNA.
 DR EMBL; Z11695; CAA7753.1; ALT_INIT; mRNA.
 DR EMBL; BC017832; AAH17832.1; -; mRNA.
 DR PIR; J01400; J01400.
 DR PDB; 1PME; X-ray; @=1-360.
 DR OGP; P28482; -.
 DR Ensembl; ENSG00000100030; Homo sapiens.
 DR HGNC; HGNC:6871; MAPK1.
 DR H-InvDB; HIX0016281; -.
 DR Reactome; P28482; -.
 DR MIM; 176948; -.
 DR GO; GO:0004707; F-MAP kinase activity; TAS.
 DR GO; GO:0006935; P:chemotaxis; TAS.
 DR GO; GO:0006917; P:induction of apoptosis; TAS.
 DR GO; GO:0006950; P:response to stress; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR GO; GO:0007268; P:synaptic transmission; TAS.
 DR InterPro; IPR008349; Erk_1_2_MAPK.
 DR InterPro; IPR003527; MAP_Kin.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR PRINTS; PR01770; ERKIERK2MAPK.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS01351; MAPK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
 DR 3D-structure; ATP-binding; Cell cycle; Kinase; Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase; Transferase.
 KW DOMAIN 25 313
 FT ATP (By similarity).
 NP_BIND 31 39
 FT Poly-Ala.
 FT COMBIAS 2 9
 FT ACT_SITE 149 149
 FT BINDING 54 54
 FT MOD_RES 185 185
 FT MOD_RES 187 187
 FT CONFLICT 91 91
 FT TURN 23 24
 FT STRAND 25 30
 FT STRAND 39 44
 FT TURN 45 47
 FT STRAND 50 56
 FT TURN 59 60
 FT TURN 62 77
 FT HELIX 81 82
 FT TURN 83 83
 FT STRAND 88 90
 FT TURN 95 97
 FT STRAND 101 106
 FT STRAND 110 111
 FT HELIX 112 118
 FT HELIX 123 142
 FT TURN 143 144
 FT STRAND 145 146
 FT STRAND 152 154
 FT HELIX 155 157
 FT STRAND 159 160
 FT TURN 163 165
 FT STRAND 169 170
 FT TURN 172 173
 FT STRAND 176 178
 FT HELIX 180 180
 FT STRAND 183 184
 FT TURN 191 193
 FT HELIX 196 198
 FT TURN 199 201

RC	STRAIN=Swiss; TISSUE=Fibroblast;	RP	SEQUENCE.
RX	MEDLINE=91305126; PubMed=1649458;	RA	MEDLINE=91184134; PubMed=1849075;
RA	Her J.-H., Wu J.-S., Kall T.B., Sturgill T.W., Weber M.J.;	RA	Payne D.M., Rossomando A.J., Martino P., Erickson A.K., Her J.-H.,
RT	"Sequence of pp42/MAP kinase, a serine/threonine kinase regulated by	RA	Shabanowitz J., Hunt D.F., Weber M.J., Sturgill T.W.;
RT	tyrosine phosphorylation."	RT	"Identification of the regulatory phosphorylation sites in
RL	Nucleic Acids Res. 19:3743-3743 (1991).	RT	pp42/mitogen-activated protein kinase (MAP kinase).";
RN	[2]	RL	EMBO J. 10:885-892 (1991).
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].	CC	FUNCTION: Involved in both the initiation and regulation of
RC	STRAIN=C57BL/6J, and NOD; TISSUE=Head, Thymus, and Urinary bladder;	CC	meiosis, mitosis, and postmitotic functions in differentiated
RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;	CC	cells by phosphorylating a number of transcription factors such as
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,	CC	ELK-1. Phosphorylates E1F4BPI; required for initiation of
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,	CC	translation. Phosphorylates microtubule-associated protein 2
RA	Vagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,	CC	(MAP2) (By similarity).
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,	CC	MAP2 (By similarity).
RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,	CC	MAP2 (By similarity).
RA	Blake J.A., Bratt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,	CC	MAP2 (By similarity).
RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,	CC	MAP2 (By similarity).
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,	CC	MAP2 (By similarity).
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,	CC	MAP2 (By similarity).
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,	CC	MAP2 (By similarity).
RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,	CC	MAP2 (By similarity).
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,	CC	MAP2 (By similarity).
RA	Kagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,	CC	MAP2 (By similarity).
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,	CC	MAP2 (By similarity).
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,	CC	MAP2 (By similarity).
RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,	CC	MAP2 (By similarity).
RA	Valtando R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,	CC	MAP2 (By similarity).
RA	Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,	CC	MAP2 (By similarity).
RA	Wu Ming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,	CC	MAP2 (By similarity).
RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,	CC	MAP2 (By similarity).
RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,	CC	MAP2 (By similarity).
RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,	CC	MAP2 (By similarity).
RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,	CC	MAP2 (By similarity).
RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,	CC	MAP2 (By similarity).
RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,	CC	MAP2 (By similarity).
RA	Birney E., Hayashizaki Y.;	CC	MAP2 (By similarity).
RT	"Analysis of the mouse transcriptome based on functional annotation of	CC	MAP2 (By similarity).
RT	60,770 full-length cDNAs."	CC	MAP2 (By similarity).
RT	Nature 420:563-573 (2002).	CC	MAP2 (By similarity).
RN	[3]	CC	MAP2 (By similarity).
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].	CC	MAP2 (By similarity).
RC	TISSUE=Eye;	CC	MAP2 (By similarity).
RX	MEDLINE=22388257; PubMed=1477932; DOI=10.1073/pnas.242603899;	CC	MAP2 (By similarity).
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	CC	MAP2 (By similarity).
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	CC	MAP2 (By similarity).
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	CC	MAP2 (By similarity).
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	CC	MAP2 (By similarity).
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	CC	MAP2 (By similarity).
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	CC	MAP2 (By similarity).
RA	Brownstein M.J., Uudin T.B., Toshiyuki S., Carninci P., Prange C.,	CC	MAP2 (By similarity).
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,	CC	MAP2 (By similarity).
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	CC	MAP2 (By similarity).
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	CC	MAP2 (By similarity).
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	CC	MAP2 (By similarity).
RA	Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,	CC	MAP2 (By similarity).
RA	Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,	CC	MAP2 (By similarity).
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	CC	MAP2 (By similarity).
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	CC	MAP2 (By similarity).
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,	CC	MAP2 (By similarity).
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;	CC	MAP2 (By similarity).
RT	"Generation and initial analysis of more than 15,000 full-length human	CC	MAP2 (By similarity).
RT	and mouse cDNA sequences."	CC	MAP2 (By similarity).
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	CC	MAP2 (By similarity).
RN	[4]	CC	MAP2 (By similarity).
RP	NUCLEOTIDE SEQUENCE OF 151-189.	CC	MAP2 (By similarity).
RC	STRAIN=CBA; TISSUE=Bone marrow;	CC	MAP2 (By similarity).
RX	MEDLINE=93185941; PubMed=8444355; DOI=10.1016/0378-1119(93)90411-U;	CC	MAP2 (By similarity).
RA	Ershler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;	CC	MAP2 (By similarity).
RT	"Novel CDC2-related protein kinases produced in murine hematopoietic	CC	MAP2 (By similarity).
RT	stem cells."	CC	MAP2 (By similarity).
RL	Gene 124:305-306 (1993).	CC	MAP2 (By similarity).
RN	[5]	CC	MAP2 (By similarity).
RP	PHOSPHORYLATION SITES THR-183 AND TYR-185, AND PARTIAL PROTEIN	CC	MAP2 (By similarity).

Query Match		81.4%;	Score 1434.5;	DB 1;	Length 358;
Best Local Similarity		73.8%;	Pred. No. 5.7e-82;		
Matches 276;		Conservative	19;	Mismatches 16;	Indels 63;
				Gaps	2;
QY	1	MAAAAGCGGGGPRTEGPGVGEVENVKQPDGPRYTQLOVIGGAGCMYSSAY 60			
DB	1	MAAAAAAG-----PEMVRGQGVDFDPRYTNSLYIGGAGCMYCSAY 41			
QY	61	DHYKTRVAIKKISPPHQTQYTCORTLREIQILFRFHNVIIGDILRASTLEAMRDVYI 120			
DB	42	DNLKRVVAIKKISPPHQTQYTCORTLREIKILFRFHNVIIGDILRASTLEAMRDVYI 101			
QY	121	VDLMETDLYKLLKSSQSLNDHICFYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180			
DB	102	VDLMETDLYKLLKSSQSLNDHICFYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 161			
QY	181	KICDFGLARADDEHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 240			
DB	162	KICDFGLARADDEHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 221			
QY	241	NRPIFGKHYLDOLNHTL----- 258			
DB	222	NRPIFGKHYLDOLNHTL----- 281			
QY	259	--ALDLDRLMTFNPKRITVEEALAHPLYEQYDPTDFVABEPPTFAMELDDLPKRL 316			
DB	282	SKALDLDRLMTFNPKRITVEEALAHPLYEQYDPTDFVABEPPTFAMELDDLPKRL 341			
QY	317	KELIFQTARFQPG 330			
DB	342	KELIFQTARFQPG 355			

RESULT 15

MK01	RAT	STANDARD;	PRT;	358 AA.
ID	P63086;	P27703;		
AC	01-AUG-1992	(Rel. 23, Created)		
DT	01-AUG-1992	(Rel. 23, Last sequence update)		
DT	13-SEP-2005	(Rel. 48, Last annotation update)		
DE	Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular			
DE	signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase			
DE	2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERT1).			
GN	Name=Mapk1; Synonyms=Erk2, Mapk, Prkml;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC	Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	Strain=Sprague-Dawley; Tissue=Brain;			
RX	MEDLINE=91235302; PubMed=2032290; DOI=10.1016/0092-8674(91)90098-J;			
RA	Boulton T.G., Nye S.H., Robbins D.J., Ip N.Y., Radziejewska E.,			
RA	Morgenbesser S.D., DePinho R.A., Panayotatos N., Cobb M.H.,			
RA	Yancopoulos G.D.;			
RT	"ERKs: a family of protein-serine/threonine kinases that are activated			
RT	and tyrosine phosphorylated in response to insulin and NGF.";			
RL	Cell 65:663-675(1991).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).			
RX	MEDLINE=94150699; PubMed=8107865; DOI=10.1038/367704a0;			
RA	Zhang F., Strand A., Robbins D., Cobb M.H., Goldsmith E.J.;			
RT	"Atomic structure of the MAP kinase ERK2 at 2.3-A resolution.";			
RL	Nature 367:704-710(1994).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).			
RX	MEDLINE=97442349; PubMed=9298998; DOI=10.1016/S0092-8674(00)80351-7;			
RA	Canagarajah B.J., Khokhlatchev A., Cobb M.H., Goldsmith E.J.;			
RT	"Activation mechanism of the MAP kinase ERK2 by dual			
RT	phosphorylation.";			
RL	Cell 90:859-869(1997).			

[4]	AUTOPHOSPHORYLATION.		
RX	MEDLINE=91296777; PubMed=1712480;		
RA	Seger R., Ahn N.G., Boulton T.G., Yancopoulos G.D., Panayotatos N.,		
RA	Radziejewska E., Ericsson L., Bratlien R.L., Cobb M.H., Krebs E.G.;		
RT	"Microtubule-associated protein 2 kinases, ERK1 and ERK2, undergo		
RT	autophosphorylation on both tyrosine and threonine residues:		
RT	implications for their mechanism of activation.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 88:6142-6146(1991).		
[5]	PHOSPHORYLATION OF EIF4EBP1.		
RX	MEDLINE=95025978; PubMed=7939721;		
RA	Lin T.-A., Kong X., Haystead T.A.J., Pause A., Belsham G.J.;		
RA	Sonenberg N., Lawrence J.C. Jr.;		
RT	"PHAS-I as a link between mitogen-activated protein kinase and		
RT	translation initiation.";		
RL	Science 266:653-656(1994).		
CC	-I- FUNCTION: Involved in both the initiation and regulation of		
CC	meiosis, mitosis, and postmitotic functions in differentiated		
CC	cells by phosphorylating a number of transcription factors such as		
CC	ELK-1. Phosphorylates EIF4EBP1; required for initiation of		
CC	translation. Phosphorylates microtubule-associated protein 2		
CC	(MAP2).		
CC	-I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.		
CC	-I- COFACTOR: Magnesium (By similarity).		
CC	-I- ENZYME REGULATION: Activated by phosphorylation on tyrosine and		
CC	threonine in response to insulin and NGF. Both phosphorylations		
CC	are required for activity.		
CC	-I- TISSUE SPECIFICITY: Highest levels within the nervous system,		
CC	expressed in different tissues, mostly in muscle, thymus and		
CC	heart.		
CC	-I- DEVELOPMENTAL STAGE: Increased expression during development.		
CC	-I- PTM: Autophosphorylated on threonine and tyrosine residues, which		
CC	correlates with a slow and low level of autoactivation.		
CC	-I- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP		
CC	kinase subfamily.		
CC	-----		
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use as long as its content is in no way modified and this statement is not		
CC	removed.		
CC	-----		
DR	EMBL; M64300; AAA41124.1; -; mRNA.		
DR	PIR; A40033; A40033.		
DR	PDB; 1ERK; X-ray; @=1-358.		
DR	PDB; 1GOL; X-ray; @=1-358.		
DR	PDB; 2ERK; X-ray; @=1-358.		
DR	PDB; 3ERK; X-ray; @=1-358.		
DR	PDB; 4ERK; X-ray; @=1-358.		
DR	SMR; P63086; 14-357.		
DR	Ensembl; ENSRNOG0000001849; Rattus norvegicus.		
DR	Reactome; P63086; -.		
DR	RGD; 70500; Mapk1.		
DR	InterPro; IPR008349; Erk_1_2_MAPK.		
DR	InterPro; IPR003527; MAP_kin.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR008271; Ser_thr_pkin_AS.		
DR	Pfam; PF00069; Pkinase; 1.		
DR	PRINTS; PR01770; ERK1ERK2MAPK.		
DR	PROSITE; PS01351; MAPK; 1.		
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.		
DR	PROSITE; PS00107; PROTEIN KINASE DOM; 1.		
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.		
KW	3D-structure; ATP-binding; Cell Cycle; Kinase; Nucleotide-binding;		
KW	Phosphorylation; Serine/threonine-protein kinase; Transferase.		
FT	DOMAIN	23	311
FT	NP_BIND	29	37
FT	COMPBIAS	2	7
FT	ACT_SITE	147	147
FT	BINDING	52	52
FT	MOD_RES	183	183
FT	ATP.		
FT	Poly-Ala.		
FT	ATP.		
FT	Phosphothreonine (activates the kinase)		
FT	(By similarity).		

Phosphotyrosine (activates the kinase) (By similarity).		185	185
FT	MOD_RES	11	12
FT	TURN	13	14
FT	STRAND	15	16
FT	TURN	20	22
FT	STRAND	23	24
FT	STRAND	28	31
FT	STRAND	35	42
FT	TURN	43	46
FT	STRAND	47	54
FT	TURN	57	58
FT	HELIX	60	75
FT	TURN	79	80
FT	STRAND	81	81
FT	STRAND	84	89
FT	TURN	93	95
FT	STRAND	99	104
FT	STRAND	108	109
FT	HELIX	110	114
FT	TURN	115	116
FT	HELIX	121	140
FT	TURN	141	142
FT	STRAND	143	144
FT	HELIX	150	152
FT	STRAND	153	155
FT	TURN	157	158
FT	STRAND	161	163
FT	TURN	167	168
FT	STRAND	170	171
FT	HELIX	174	176
FT	STRAND	178	178
FT	TURN	181	182
FT	HELIX	189	191
FT	HELIX	194	196
FT	TURN	197	199
FT	STRAND	200	200
FT	TURN	205	205
FT	HELIX	206	221
FT	HELIX	231	242
FT	HELIX	247	252
FT	HELIX	256	264
FT	HELIX	273	276
FT	TURN	278	279
FT	HELIX	282	291
FT	HELIX	296	298
FT	HELIX	302	306
FT	TURN	307	307
FT	HELIX	309	311
FT	TURN	312	314
FT	HELIX	317	319
FT	TURN	330	331
FT	HELIX	338	348
FT	TURN	349	349
FT	HELIX	350	352
SQ	SEQUENCE	358 AA;	41276 MW; 3BBCF22471EDBA0B CRC64;

Query Match 81.4%; Score 1434.5; DB 1; Length 358;
Best Local Similarity 73.8%; Pred. No. 5.7e-82;
Matches 276; Conservative 19; Mismatches 16; Indels 63; Gaps 2;
QY 1 MAAAAAGGGGPRTEGVGPGEVEMVKGQPFDPVPTQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAG-----PENVRQVDFVDPRTNLSYIGEGAYGMVCSAY 41
QY 61 DHVKTRVAIKKISPEHQYQCORTLREIQILFRPHENVIGIRDILRASTLEAMRDVYI 120
Db 42 DNLNKRVAIKKISPEHQYQCORTLREIKILFRPHENIIGINDIIRAPTIEQMKDVI 101
QY 121 VQDLMETDLYKLIKSQLSDNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
Db 102 VQDLMETDLYKLLKTLQHLSDNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 161

Search completed: February 6, 2006, 15:34:38
Job time : 164.777 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:34:56 ; Search time 33.7896 Seconds
(without alignments)
819.670 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAAQQGGGGPRRTG.....LKELIFQETARFQPGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1731	98.2	379	2	US-09-411-628-14
2	1731	98.2	379	2	US-10-174-794-14
3	1731	98.2	379	2	US-09-538-092-1021
4	1725	97.8	379	2	US-09-025-580-25
5	1725	97.8	379	2	US-09-642-749-25
6	1725	97.8	631	2	US-09-417-197-39
7	1719	97.5	379	2	US-08-622-277A-8
8	1659	94.1	624	2	US-09-417-197-57
9	1607	91.2	367	1	US-08-176-620A-11
10	1606	91.1	355	2	US-08-461-985-11
11	1606	91.1	355	2	US-08-459-953A-8
12	1595	90.5	365	1	US-09-393-212-8
13	1595	90.5	365	1	US-08-176-620A-2
14	1595	90.5	365	1	US-08-463-862-2
15	1595	90.5	365	1	US-08-461-985-2
16	1595	90.5	365	1	US-08-458-887-2
17	1595	90.5	365	2	US-08-932-787B-2
18	1595	90.5	365	2	US-08-932-012C-2
19	1595	90.5	365	2	US-08-888-818C-2
20	1459.5	82.8	380	2	US-09-949-016-9919
21	1454.5	82.5	380	2	US-08-459-953A-9
22	1454.5	82.5	380	2	US-09-393-212-9
23	1444.5	81.9	360	2	US-09-457-040B-3
24	1444.5	81.9	360	2	US-09-538-092-1028
25	1436	81.5	415	1	US-08-176-620A-4
26	1436	81.5	415	1	US-08-463-862-4
27	1436	81.5	415	1	US-08-461-985-4

28	1436	81.5	415	1	US-08-458-887-4	Sequence 4, Appli
29	1436	81.5	415	2	US-08-932-787B-4	Sequence 4, Appli
30	1436	81.5	415	2	US-08-932-012C-4	Sequence 4, Appli
31	1436	81.5	415	2	US-08-888-818C-4	Sequence 4, Appli
32	1435.5	81.4	360	2	US-08-622-277A-14	Sequence 14, Appli
33	1434.5	81.4	358	1	US-08-176-620A-12	Sequence 12, Appli
34	1434.5	81.4	358	1	US-08-461-985-12	Sequence 12, Appli
35	1434.5	81.4	358	2	US-09-457-040B-4	Sequence 4, Appli
36	1434.5	81.4	364	2	US-09-457-040B-34	Sequence 34, Appli
37	1434.5	81.4	604	2	US-09-417-197-59	Sequence 59, Appli
38	1434.5	81.4	605	2	US-09-417-197-41	Sequence 41, Appli
39	1421.5	80.6	365	2	US-09-457-040B-35	Sequence 35, Appli
40	1415.5	80.3	360	2	US-09-025-580-2	Sequence 2, Appli
41	1415.5	80.3	360	2	US-09-642-749-2	Sequence 2, Appli
42	905	51.3	183	1	US-08-176-620A-8	Sequence 8, Appli
43	905	51.3	183	1	US-08-461-985-8	Sequence 8, Appli
44	905	51.3	183	2	US-08-932-787B-16	Sequence 16, Appli
45	905	51.3	183	2	US-08-932-012C-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-09-411-628-14

; Sequence 14, Application US/09411628

; Patent No. 6428994

; GENERAL INFORMATION:

; APPLICANT: University of Southern California

; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN

; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES

; FILE REFERENCE: 13761-707

; CURRENT APPLICATION NUMBER: US/09/411,628

; CURRENT FILING DATE: 1999-10-01

; EARLIER APPLICATION NUMBER: US 60/102,906

; EARLIER FILING DATE: 1998-10-02

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14

; LENGTH: 379

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-411-628-14

Query Match	98.2%;	Score 1731;	DB 2;	Length 379;
Best Local Similarity	88.4%;	Pred. No. 1.2e-173;		
Matches	335;	Conservative	0;	Mismatches 44; Gaps 1;
Qy	1	MAAAAAQGGGGPRRTG	VGPGVGEVEMVKQPF	DVGPRYTQIQYIGGAYGMVSSAY 60
Db	1	MAAAAAQGGGGPRRTG	VGPGVGEVEMVKQPF	DVGPRYTQIQYIGGAYGMVSSAY 60
Qy	61	DHVRKTRVAIKKIS	PFPHQTYCORTREI	QILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db	61	DHVRKTRVAIKKIS	PFPHQTYCORTREI	QILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Qy	121	VQDLMETDLYKLKKSQ	LSNDHICYFLYQIL	RGKLYHSAANVLRDLKPSNLLINTTCDL 180
Db	121	VQDLMETDLYKLKKSQ	LSNDHICYFLYQIL	RGKLYHSAANVLRDLKPSNLLINTTCDL 180
Qy	181	KICDFGLARIADPEH	DHTGFLTYVATRW	TRAPFIMLSNGYTKSIDWSVGCILAEMLS 240
Db	181	KICDFGLARIADPEH	DHTGFLTYVATRW	TRAPFIMLSNGYTKSIDWSVGCILAEMLS 240
Qy	241	NRPIFPCKHYLDQNL	HL-----	----- 258
Db	241	NRPIFPCKHYLDQNL	HL-----	----- 258
Qy	259	--ALDILLDRMLTF	PNKRIITVEEALAH	PYLEQYDPTDEPVABEPTTFAMELDLPKRL 316
Db	301	SKALDILLDRMLTF	PNKRIITVEEALAH	PYLEQYDPTDEPVABEPTTFAMELDLPKRL 360
Qy	317	KELIFQETARFQPG	VLEAP 335	

```

Db      361 KELIFQETARFQGVLEAP 379
|||||
RESULT 2
US-10-174-794-14
; Sequence 14, Application US/10174794
; Patent No. 6664086
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; TITLE OF INVENTION: SEQUENCES OF LEARNING-INDUCED KINASES
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-174-794-14

Query Match      98.2%; Score 1731; DB 2; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.2e-173;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY      1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGQPFDPVGRYTOQYIGEGAYGMVSSAY 60
Db      1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGQPFDPVGRYTOQYIGEGAYGMVSSAY 60
|||||
QY      61 DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db      61 DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
|||||
QY      121 VDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db      121 VDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
|||||
QY      181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db      181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
|||||
QY      241 NRPIFGKHYLDQNLNHL----- 258
Db      241 NRPIFGKHYLDQNLNHL----- 258
|||||
QY      259 --ALDLDRLMTNPNKRITVEALAHPLYEQYDPTDEPVAEPTFAMELDDLPKERL 316
Db      301 SKALDLDRLMTNPNKRITVEALAHPLYEQYDPTDEPVAEPTFAMELDDLPKERL 360
|||||
QY      317 KELIFQETARFQGVLEAP 335
Db      361 KELIFQETARFQGVLEAP 379
|||||

RESULT 3
US-09-538-092-1021
; Sequence 1021, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01

```

```

; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurPatSeqFormatter Version 0.9
; SEQ ID NO 1021
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P27361
US-09-538-092-1021

Query Match      98.2%; Score 1731; DB 2; Length 379;
Best Local Similarity 88.4%; Pred. No. 1.2e-173;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY      1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGQPFDPVGRYTOQYIGEGAYGMVSSAY 60
Db      1 MAAAAAGGGGGPRRTGEGVPGVGEVEMVKGQPFDPVGRYTOQYIGEGAYGMVSSAY 60
|||||
QY      61 DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db      61 DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
|||||
QY      121 VDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db      121 VDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180
|||||
QY      181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db      181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
|||||
QY      241 NRPIFGKHYLDQNLNHL----- 258
Db      241 NRPIFGKHYLDQNLNHL----- 258
|||||
QY      259 --ALDLDRLMTNPNKRITVEALAHPLYEQYDPTDEPVAEPTFAMELDDLPKERL 316
Db      301 SKALDLDRLMTNPNKRITVEALAHPLYEQYDPTDEPVAEPTFAMELDDLPKERL 360
|||||
QY      317 KELIFQETARFQGVLEAP 335
Db      361 KELIFQETARFQGVLEAP 379
|||||

RESULT 4
US-09-025-580-25
; Sequence 25, Application US/09025580
; Patent No. 6162613
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; APPLICANT: Fox, Ted
; APPLICANT: Willson, Keith Phillip
; APPLICANT: Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; TITLE OF INVENTION: Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580

```

```
;
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-580-25

Query Match          97.8%; Score 1725; DB 2; Length 379;
Best Local Similarity 88.1%; Pred. No. 5.1e-173;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRTEGPGVPGVEVMVKGPFDVGPRTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGGGGGPRTEGPGVPGVEVMVKGPFDVGPRTQLOVIGEGAYGMVSSAY 60
QY 61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMTDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMTDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFPKHYLDLQNLHIL----- 258
DB 241 NRPIFPKHYLDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLPPKSD 300
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPFTFAMELDDLPKERL 316
DB 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPFTFAMELDDLPKERL 360
QY 317 KELIFQTARFQGVLEAP 335
DB 361 KELIFQTARFQGVLEAP 379

RESULT 5
US-09-642-749-25
; Sequence 25, Application US/09642749
; Patent No. 6849716
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; Fox, Ted
; Wilson, Keith Phillip
; Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,749
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
; US-09-642-749-25
```

```
Query Match          97.8%; Score 1725; DB 2; Length 379;
Best Local Similarity 88.1%; Pred. No. 5.1e-173;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRTEGPGVPGVEVMVKGPFDVGPRTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGGGGGPRTEGPGVPGVEVMVKGPFDVGPRTQLOVIGEGAYGMVSSAY 60
QY 61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMTDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMTDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFPKHYLDLQNLHIL----- 258
DB 241 NRPIFPKHYLDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLPPKSD 300
QY 259 --ALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPFTFAMELDDLPKERL 316
DB 301 SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYYDPTDEPVASEPFTFAMELDDLPKERL 360
QY 317 KELIFQTARFQGVLEAP 335
DB 361 KELIFQTARFQGVLEAP 379
```

```
RESULT 6
US-09-417-197-39
; Sequence 39, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
; FILE REFERENCE: 3759-0110P
; CURRENT FILING DATE: 1999-10-07
; CURRENT APPLICATION NUMBER: US/09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 39
```

```

; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Erk1 fusion construct
US-09-417-197-39

Query Match          97.8%; Score 1725; DB 2; Length 631;
Best Local Similarity 88.1%; Pred. No. 1.1e-172;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 60
Db 253 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 312
QY 61 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 313 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 372
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 373 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 432
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
Db 433 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 492
QY 241 NRPIPGKHLYDQNLNHL----- 258
Db 493 NRPIPGKHLYDQNLNHLILGILGSPSQEDLNCIINKARNYLSQPSKTKVAVAKLFPKSD 552
QY 259 --ALDLDRLMTFNPKNRITVVEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 316
Db 553 SKALDLDRLMTFNPKNRITVVEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 612
QY 317 KELIFQETARFQGVLEAP 335
Db 613 KELIFQETARFQGVLEAP 631

RESULT 7
US-08-622-277A-8
; Sequence 8, Application US/08622277A
; Patent No. 6001580
; GENERAL INFORMATION:
; APPLICANT: Tani, Akiyoshi
; APPLICANT: Ichimori, Yuzo
; TITLE OF INVENTION: Method For Assaying MAP Kinase
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,277A
; FILING DATE: 27-MAR-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 320577-1995
; FILING DATE: 08-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 305456-1995
; FILING DATE: 24-NOV-1995
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: JP 070125-1995
; FILING DATE: 28-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: toohey, Kimberlin M.
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 04221.0039-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
US-08-622-277A-8

Query Match          97.5%; Score 1719; DB 2; Length 379;
Best Local Similarity 87.9%; Pred. No. 2.2e-172;
Matches 333; Conservative 0; Mismatches 2; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGEPRTGEGVGPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIPGKHLYDQNLNHL----- 258
Db 241 NRPIPGKHLYDQNLNHLILGILGSPSQEDLNCIINKARNYLSQPSKTKVAVAKLFPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVVEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 316
Db 301 SKALDLDRLMTFNPKNRITVVEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 360
QY 317 KELIFQETARFQGVLEAP 335
Db 361 KELIFQETARFQGVLEAP 379

RESULT 8
US-09-417-197-57
; Sequence 57, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mERK1-F64L-S65T-GFP fusion
US-09-417-197-57

Query Match          94.1%; Score 1659; DB 2; Length 624;

```


Best Local Similarity	85.48;	Pred. No. 9.7e-166;	
Matches 323;	Conservative 1;	Mismatches 10;	Indels 44; Gaps 1;
QY	2	AAAAAGGGGGEPRTTGGVPGVGVEMVKGPQFDVGPRTYQLQYIGEGAYGMVSSAYD	61
DB	3	AAAAAPGGGGEPRTAGVVPVPGVEVVKGPQFDVGPRTYQLQYIGEGAYGMVSSAYD	62
QY	62	HVKTRVAIKKISPFHEQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVIV	121
DB	63	HVKTRVAIKKISPFHEQTYCQRTLREIQILLRFRHENVIGIRDLRAPTLEAMRDVIV	122
QY	122	QDLMETDLYKLLKSQQLSNDHI CYFLYQLIRGLKXYIHSANVLRDLKPSNLLINTTCOLK	181
DB	123	QDLMETDLYKLLKSQQLSNDHI CYFLYQLIRGLKXYIHSANVLRDLKPSNLLINTTCOLK	182
QY	182	ICDFGLARIADPEHDHDTGFTVEYVATRWYRAPEITMLNSKGYTKSIDIWSVGCILAEMLSN	241
DB	183	ICDFGLARIADPEHDHDTGFTVEYVATRWYRAPEITMLNSKGYTKSIDIWSVGCILAEMLSN	242
QY	242	RPFPFGKHYLDQLNHIL-----	258
DB	243	RPFPFGKHYLDQLNHILGILGSPQEDLNCIINMKARNYLSQPSKTKVAKLPPKSDS	302
QY	259	-ALDLDRLMTFNPNKRTTVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERLK	317
DB	303	KALDLDRLMTFNPNKRTTVEEALAHPLYEQYDPTDEPVAEEPTTFDMELODDLPKERLK	362
QY	318	ELIFQETARFQPGVLEAP	335
DB	363	ELIFOETARFQPGAPEGF	380

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

```

;          TOPOLOGY: unknown
;          MOLECULE TYPE: protein
US-08-176-620A-11

Query Match          91.2%; Score 1607; DB 1; Length 367;
Best Local Similarity 85.3%; Pred. No. 1.3e-150;
Matches 313; Conservative 1; Mismatches 9; Indels 44; Gaps 1

QY      13  EPRTEGVGPGVEVMKGGQPPDVGPRTYQLOVIGEGAYGMVSSAYDHYRKRTRVAIKK 72
      1  EPGGTAGVWPVFGEEVWVGQPPDVGPRTYQLOVIGEGAYGMVSSAYDHYRKRTRVAIKK 60
QY      73  ISPFHQYTQCTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKL 132
      61  ISPFHQYTQCTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKL 120
QY      133 LKSQLSNDHICVFLYQIILRGLKYIHANVLHRLDKPSNLLINTTCDLKI CDGGLARIAD 192
      121 LKSQLSNDHICVFLYQIILRGLKYIHANVLHRLDKPSNLLINTTCDLKI CDGGLARIAD 180
QY      193 PEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCI LAEMLSNRRPIFGPKHYLD 252
      181 PEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCI LAEMLSNRRPIFGPKHYLD 240
QY      253 QLNHL-----ALDLDRLMT 268
      241 QLNHLGILGSPSOEDLNCIINMKARYLQSLPKTKVAAKLFPKSDSKALDLDRLMT 300
QY      269 FPNKRTIVFEALAHPLYEQYDPTDPSVAEPEPTFAMELDDLPERLKELI FQETARFQ 328
      301 FPNKRTIVFEALAHPLYEQYDPTDPSVAEPEPTFAMELDDLPERLKELI FQETARFQ 360
QY      329 PGVLEAP 335
      361 PGAEAP 367

RESULT 10
US-08-461-985-11
; Sequence 11, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:

```

```
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 367 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ US-08-461-985-11

Query Match 91.2%; Score 1607; DB 1; Length 367;
Best Local Similarity 85.3%; Pred. No. 1.3e-160;
Matches 313; Conservative 1; Mismatches 9; Indels 44; Gaps 1;

QY 13 EPRRTGVGPGVGEVEMVKGQPFVDPRTYQYIGEGAGVSSAYDHRVTRVAIKK 72
DB 1 EPRGTAGVVPVGEVEVVKGPFDVGPRTYQYIGEGAGVSSAYDHRVTRVAIKK 60
QY 73 ISPFHQTYCORTLRIQIILLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKL 132
DB 61 ISPFHQTYCORTLRIQIILLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKL 120
QY 133 LKSQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIAD 192
DB 121 LKSQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIAD 180
QY 193 PEHDTGFTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLD 252
DB 181 PEHDTGFTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLD 240
QY 253 QLNHIL-----ALDLDRLMT 268
DB 241 QLNHILGILGSPSQEDLNCIINNKARNYLOSLPSKTKVAVAKLFPKSDSKALDLDRLMT 300
QY 269 FNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQ 328
DB 301 FNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQ 360
QY 329 PGVLEAP 335
DB 361 PGVLEAP 367

RESULT 11
US-08-459-953A-8
; Sequence 8, Application US/08459953A
; Patent No. 6030822
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:

/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 367 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ US-08-461-985-11

Query Match 91.1%; Score 1606; DB 2; Length 355;
Best Local Similarity 87.6%; Pred. No. 1.6e-160;
Matches 311; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 25 PGEVEMVKGQPFVDPRTYQYIGEGAGVSSAYDHRVTRVAIKKISPEHQTYCOR 84
DB 1 PGEVEMVKGQPFVDPRTYQYIGEGAGVSSAYDHRVTRVAIKKISPEHQTYCOR 60
QY 85 TLREIQILRLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLKSQLSNDHIC 144
DB 61 TLREIQILRLFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLKSQLSNDHIC 120
QY 145 YFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFTLEY 204
DB 121 YFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFTLEY 180
QY 205 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQNLHIL----- 258
DB 181 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIFPKGHYLDQNLHILGILGSP 240
QY 259 -----ALDLDRLMTFNPKNRITVEEA 280
DB 241 SQEDLNCIINNKARNYLOSLPSKTKVAVAKLFPKSDSKALDLDRLMTFNPKNRITVEEA 300
QY 281 LAHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQGVLEAP 335
DB 301 LAHPYLEQYDPTDEPVAEPPFTFAMELDDLPKRLKELIFQETARFQGVLEAP 355

RESULT 12
US-09-393-212-8
; Sequence 8, Application US/09393212
; Patent No. 6579972
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/393,212
FILING DATE: 09-Sep-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/459,953A
FILING DATE: June 2, 1995
APPLICATION NUMBER: 08/029,494
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/267
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-393-212-8

Query Match 91.1%; Score 1606; DB 2; Length 355;
Best Local Similarity 87.6%; Pred. No. 1.6e-160;
Matches 311; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 25 PGEVWVGQPPDVGPRYTQLOVIGEGAYGWSAYDHVKTRVAIKKISPPFHTQYCOR 84
DB 1 PGEVWVGQPPDVGPRYTQLOVIGEGAYGWSAYDHVKTRVAIKKISPPFHTQYCOR 60
QY 85 TLREIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKSSQSLNDHIC 144
DB 61 TLREIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLKSSQSLNDHIC 120
QY 145 YFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFLTEY 204
DB 121 YFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDTGFLTEY 180
QY 205 VATRWYRAPIMLNSKGYTKSIDWSVGCILAEMLSNRPIFPKGYLDQLNHIL----- 258
DB 181 VATRWYRAPIMLNSKGYTKSIDWSVGCILAEMLSNRPIFPKGYLDQLNHILGILGSP 240
QY 259 -----ALDLDRLMTFNPKNRITVEEA 280
DB 241 SQEDLNCIINMKARNYLOSLPSKTKVAKLFPKSDSKALDLDRLMTFNPKNRITVEEA 300
QY 281 LAHPYLEQYDPTDEPVAEPPTFAMELDLDPKRLKELIFQETARFPQGVLEAP 335
DB 301 LAHPYLEQYDPTDEPVAEPPTFAMELDLDPKRLKELIFQETARFPQGVLEAP 355

RESULT 13
US-08-176-620A-2
Sequence 2, Application US/08176620A
Patent No. 5595904
GENERAL INFORMATION:
APPLICANT: Boulton, Teri G.
APPLICANT: Cobb, Melanie H.
APPLICANT: Yancopoulos, George D.
APPLICANT: Nye, Steven
APPLICANT: Panayotatos, Nikos
TITLE OF INVENTION: A Family of Map2 Protein Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,620A
FILING DATE: 03-JAN-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Lealie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-620A-2

Query Match 90.5%; Score 1595; DB 1; Length 365;
Best Local Similarity 85.2%; Pred. No. 2.4e-159;
Matches 311; Conservative 1; Mismatches 9; Indels 44; Gaps 1;
QY 15 RTEGVGPGVGEVEMVKGQPPDVGPRYTQLOVIGEGAYGWSAYDHVKTRVAIKKIS 74
DB 1 RGTAGVVPVPGVEVEMVKGQPPDVGPRYTQLOVIGEGAYGWSAYDHVKTRVAIKKIS 60
QY 75 PFEHQTQYCORLEIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLK 134
DB 61 PFEHQTQYCORLEIQILLFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLK 120
QY 135 SQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPE 194
DB 121 SQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPE 180
QY 195 HDHTGFLTEVATRWYRAPIMLNSKGYTKSIDWSVGCILAEMLSNRPIFPKGYLDQL 254
DB 181 HDHTGFLTEVATRWYRAPIMLNSKGYTKSIDWSVGCILAEMLSNRPIFPKGYLDQL 240
QY 255 NHIL-----ALDLDRLMTFN 270
DB 241 NHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAKLFPKSDSKALDLDRLMTFN 300
QY 271 PNKRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDLDPKRLKELIFQETARFPQ 330
DB 301 PNKRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDLDPKRLKELIFQETARFPQ 360
QY 331 VLEAP 335
DB 361 APEAP 365

RESULT 14
US-08-463-862-2
Sequence 2, Application US/08463862
Patent No. 5776751
GENERAL INFORMATION:
APPLICANT: Boulton, Teri G. et al.
TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

Search completed: February 6, 2006, 15:36:48
Job time : 34.7896 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:35:46 ; Search time 116.816 Seconds
(without alignments)
1198.238 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAQGGGGPRTEGV.....LKELIFQETARFQPGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	335	5	US-10-623-108-2
2	1731	98.2	379	4	US-10-233-448-7
3	1731	98.2	379	4	US-10-174-794-14
4	1731	98.2	379	4	US-10-394-322A-19
5	1731	98.2	379	4	US-10-701-490-8
6	1731	98.2	379	5	US-10-735-118-3
7	1731	98.2	379	5	US-10-623-108-6
8	1731	98.2	379	5	US-10-945-684-2
9	1725	97.8	379	4	US-10-060-065-16
10	1725	97.8	379	4	US-10-059-585-37
11	1725	97.8	379	4	US-10-170-663-25
12	1725	97.8	379	5	US-10-893-072-25
13	1725	97.8	631	4	US-10-072-036-39
14	1659	94.1	624	4	US-10-072-036-57
15	1606	91.1	355	4	US-10-461-402-8
16	1599	90.7	359	5	US-10-623-108-4
17	1599	90.7	359	5	US-10-623-108-8
18	1593	90.4	359	4	US-10-114-270-110
19	1455.5	82.6	392	5	US-10-732-923-1420
20	1454.5	82.5	380	4	US-10-461-402-9
21	1444.5	81.9	360	4	US-10-171-311-127
22	1444.5	81.9	360	4	US-10-233-448-8
23	1444.5	81.9	360	4	US-10-301-822-115
24	1444.5	81.9	360	4	US-10-394-322A-20
25	1444.5	81.9	360	5	US-10-941-635-56
26	1438.5	81.6	360	4	US-10-664-421-56
27	1434.5	81.4	358	5	US-10-945-684-1

28	1434.5	81.4	358	5	US-10-451-291B-10	Sequence 10, Appl
29	1434.5	81.4	604	4	US-10-072-036-59	Sequence 59, Appl
30	1434.5	81.4	605	4	US-10-072-036-41	Sequence 41, Appl
31	1433	81.3	369	5	US-10-732-923-1421	Sequence 1421, Ap
32	1429	81.1	356	4	US-10-664-421-19	Sequence 19, Appl
33	1426	80.9	355	5	US-10-941-635-19	Sequence 19, Appl
34	1424	80.8	348	5	US-10-756-149-4856	Sequence 4856, Ap
35	1415.5	80.3	360	4	US-10-170-663-2	Sequence 2, Appli
36	1415.5	80.3	360	5	US-10-893-072-2	Sequence 2, Appli
37	1372	77.8	351	5	US-10-732-923-1406	Sequence 1406, Ap
38	1358	77.0	369	5	US-10-732-923-1456	Sequence 1456, Ap
39	1318.5	74.8	395	5	US-10-732-923-1424	Sequence 1424, Ap
40	1312.5	74.4	362	5	US-10-732-923-1418	Sequence 1418, Ap
41	1272	72.1	376	4	US-10-369-493-5689	Sequence 5689, Ap
42	919	52.1	353	5	US-10-732-923-1492	Sequence 1492, Ap
43	916	52.0	355	5	US-10-732-923-1361	Sequence 1361, Ap
44	916	52.0	355	5	US-10-732-923-1495	Sequence 1495, Ap
45	914	51.8	355	5	US-10-732-923-1287	Sequence 1287, Ap

ALIGNMENTS

RESULT 1

US-10-623-108-2
; Sequence 2, Application US/106233108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHMO
; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-108-2

Query Match	100.0%;	Score 1763;	DB 5;	Length 335;
Best Local Similarity	100.0%;	Pred. No. 1e-136;		
Matches 335;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAAAAQGGGGPRTEGVGPGVEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60	
Db	1	MAAAAQGGGGPRTEGVGPGVEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60	
Qy	61	DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDIRASTLEAMRDVYI	120	
Db	61	DHVRKTRVAIKKISPFHQYTCORTLREIQILLRFRHENVIGIRDIRASTLEAMRDVYI	120	
Qy	121	VDLMETDLYKLKSKQSLNDHICYFLYQILRLKYIHSANVLRDLKPSNLLINTTCDL	180	
Db	121	VDLMETDLYKLKSKQSLNDHICYFLYQILRLKYIHSANVLRDLKPSNLLINTTCDL	180	
Qy	181	KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240	
Db	181	KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240	
Qy	241	NRPIFFPKHYLDQNLHILADLLDRMLTFPNKRITVEEALAHPLYEQYDPTDPEVAEE	300	
Db	241	NRPIFFPKHYLDQNLHILADLLDRMLTFPNKRITVEEALAHPLYEQYDPTDPEVAEE	300	
Qy	301	PFTFAMELDDLPKRLKELIFQETARFQPGVLEAP	335	
Db	301	PFTFAMELDDLPKRLKELIFQETARFQPGVLEAP	335	

RESULT 2

US-10-233-448-7
; Sequence 7, Application US/10233448

```
; Publication No. US20030109419A1
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; TITLE OF INVENTION: METHOD FOR CLASSIFICATION OF ANTI-PSYCHOTIC DRUGS
; FILE REFERENCE: 11181-009
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 60/316,338
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-233-448-7

Query Match      98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRRTGEGVGPGEVEMVKGPDPVGPRTYQIQYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAQQGGGGEPRRTGEGVGPGEVEMVKGPDPVGPRTYQIQYIGEGAYGMVSSAY 60
   |||||
QY 61 DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
   |||||
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
QY 241 NRPIFGKHLYDLQNLHIL----- 258
   |||||
Db 241 NRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPPKSD 300
   |||||
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPPKERL 316
   |||||
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPPKERL 360
   |||||
QY 317 KELIFOETARFQGVLEAP 335
   |||||
Db 361 KELIFOETARFQGVLEAP 379
   |||||

RESULT 3
US-10-174-794-14
; Sequence 14, Application US/10174794
; Publication No. US20030166220A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; PRIOR FILING DATE: 2002-06-18
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-174-794-14
```

```
; Publication No. US20030109419A1
; GENERAL INFORMATION:
; APPLICANT: Filsone, Gilberto
; TITLE OF INVENTION: METHOD FOR CLASSIFICATION OF ANTI-PSYCHOTIC DRUGS
; FILE REFERENCE: 11181-009
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 60/316,338
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-233-448-7

Query Match      98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRRTGEGVGPGEVEMVKGPDPVGPRTYQIQYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAQQGGGGEPRRTGEGVGPGEVEMVKGPDPVGPRTYQIQYIGEGAYGMVSSAY 60
   |||||
QY 61 DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
   |||||
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
QY 241 NRPIFGKHLYDLQNLHIL----- 258
   |||||
Db 241 NRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPPKSD 300
   |||||
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPPKERL 316
   |||||
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPPKERL 360
   |||||
QY 317 KELIFOETARFQGVLEAP 335
   |||||
Db 361 KELIFOETARFQGVLEAP 379
   |||||

RESULT 4
US-10-394-322A-19
; Sequence 19, Application US/10394322A
; Publication No. US20030232391A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-394-322A-19

Query Match      98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRRTGEGVGPGEVEMVKGPDPVGPRTYQIQYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAQQGGGGEPRRTGEGVGPGEVEMVKGPDPVGPRTYQIQYIGEGAYGMVSSAY 60
   |||||
QY 61 DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLINTTCDL 180
   |||||
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
```


Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHIL----- 258
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVAMAKLFPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 316
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379

RESULT 5

US-10-701-490-8
; Sequence 8, Application US/10701490
; Publication No. US20040106141A1
; GENERAL INFORMATION:
; APPLICANT: PAUL S. MISCHER
; APPLICANT: CHARLES L. SAWYERS
; APPLICANT: BRADLEY L. SMITH
; APPLICANT: KATHERINE CROSBY
; TITLE OF INVENTION: METHODS AND MATERIALS FOR EXAMINING
; FILE REFERENCE: PATHWAYS ASSOCIATED WITH GLOBLASTOMA PROGRESSION
; CURRENT APPLICATION NUMBER: US/10/701,490
; CURRENT FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/423,777
; PRIOR FILING DATE: 2002-11-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-701-490-8

Query Match 98.2%; Score 1731; DB 4; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAGGGGGPRTEGVPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRTEGVPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHIL----- 258
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVAMAKLFPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 316
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379

RESULT 6

US-10-735-118-3
; Sequence 3, Application US/10735118
; Publication No. US20040248151A1
; GENERAL INFORMATION:
; APPLICANT: Bacus, Sarah S.
; APPLICANT: Smith, Bradley L.
; TITLE OF INVENTION: METHOD FOR PREDICTING THE RESPONSE TO HER2-DIRECTED THERAPY
; FILE REFERENCE: 6270-701.201
; CURRENT APPLICATION NUMBER: US/10/735,118
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 60/370,473
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/432,943
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 10/408,520
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-735-118-3

Query Match 98.2%; Score 1731; DB 5; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
QY 1 MAAAAAGGGGGPRTEGVPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRTEGVPGEVEMVKGQPDVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIVSGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHIL----- 258
Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVAMAKLFPKSD 300
QY 259 --ALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 316
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKERL 360
QY 317 KELIFOETARFQGVLEAP 335
Db 361 KELIFOETARFQGVLEAP 379

RESULT 7

US-10-623-108-6
; Sequence 6, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHWO
; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens

```
US-10-623-108-6
Query Match      98.2%; Score 1731; DB 5; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRTTEGVGPGVGEVEMVKGQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGEPRTTEGVGPGVGEVEMVKGQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEFHQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEFHQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIPPGKHYLDQLNHIL----- 258
DB 241 NRPIPPGKHYLDQLNHILGILGSPSQEDLNCIINMKARNYLSQSLPKTKVAVAKLPKSD 300
QY 259 --ALDILLRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 316
DB 301 SKALDILLRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLPKERL 360
QY 317 KELIFQETARFQGVLEAP 335
DB 361 KELIFQETARFQGVLEAP 379

RESULT 8
US-10-945-684-2
; Sequence 2, Application US/10945684
; Publication No. US2005009567A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Methods and Kits for Detecting Proteins
; FILE REFERENCE: 50508-2280
; CURRENT APPLICATION NUMBER: US/10/945,684
; CURRENT FILING DATE: 2004-09-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-945-684-2

Query Match      98.2%; Score 1731; DB 5; Length 379;
Best Local Similarity 88.4%; Pred. No. 5.2e-134;
Matches 335; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRTTEGVGPGVGEVEMVKGQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGEPRTTEGVGPGVGEVEMVKGQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEFHQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEFHQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240

US-10-060-065-16
; Sequence 16, Application US/10060065
; Publication No. US20030017480A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Negai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-Ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/10/060,065
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-065-16

Query Match      97.8%; Score 1725; DB 4; Length 379;
Best Local Similarity 88.1%; Pred. No. 1.6e-133;
Matches 334; Conservative 1; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGEPRTTEGVGPGVGEVEMVKGQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGEPRTTEGVGPGVGEVEMVKGQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEFHQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEFHQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
```

QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDINWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDINWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHIL-----258
DB 241 NRPIFGKHYLDQNLHIL-----258
QY 259 --ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRL 316
DB 301 SKALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRL 360
QY 317 KELIFOETARFQGVLEAP 335
DB 361 KELIFOETARFQGVLEAP 379
RESULT 10
US-10-059-585-37
; Sequence 37, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-Ichi
; APPLICANT: Nezu, Jun-Ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-059-585-37

Query Match 97.8%; Score 1725; DB 4; Length 379;
Best Local Similarity 88.1%; Pred. No. 1.6e-133;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;
QY 1 MAAAAACGGCGGPRRTGEGVGPVGEVEMVKGQPFDPGPRYTQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAACGGCGGPRRTGEGVGPVGEVEMVKGQPFDPGPRYTQIQYIGEGAYGMVSSAY 60
QY 61 DHVAKTRVAIKKISPFHQYTCQRTLRRIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
DB 61 DHVAKTRVAIKKISPFHQYTCQRTLRRIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120

QY 121 VQDLMETDYKLLKSQOLSNHDHCYFYQILRLGKLYIHSANVLHRLDKPSNLLINTTCDL 180
DB 121 VQDLMETDYKLLKSQOLSNHDHCYFYQILRLGKLYIHSANVLHRLDKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDINWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDINWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHIL-----258
DB 241 NRPIFGKHYLDQNLHIL-----258
QY 259 --ALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRL 316
DB 301 SKALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDEPVABEPTTFAMELDDLPKRL 360
QY 317 KELIFOETARFQGVLEAP 335
DB 361 KELIFOETARFQGVLEAP 379
RESULT 11
US-10-170-663-25
; Sequence 25, Application US/10170663
; Publication No. US20030165899A1
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; APPLICANT: Fox, Ted
; APPLICANT: Wilson, Keith Phillip
; APPLICANT: Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/170,663
; FILING DATE: 12-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-170-663-25

Query Match 97.8%; Score 1725; DB 4; Length 379;
Best Local Similarity 88.1%; Pred. No. 1.6e-133;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

```
QY 1 MAAAAQGGGGPRTEGVGPGVGEVEMVKGPFDVGPRTQLOYGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGVGPGVGEVEMVKGPFDVGPRTQLOYGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLRQIILLRFHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHQTYCQRTLRQIILLRFHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
DB 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHIL----- 258
DB 241 NRPIFGKHYLDQNLHILGLGSPSQEDLNCIINMKARNVLQSLPSKTKVAVAKLFPKSD 300
QY 259 --ALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLKPERL 316
DB 301 SKALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLKPERL 360
QY 317 KELIFOETARFQGVLEAP 335
DB 361 KELIFOETARFQGVLEAP 379
```

```
RESULT 12
US-10-893-072-25
; Sequence 25, Application US/10893072
; Publication No. US20040259166A1
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; Fox, Ted
; Wilson, Keith Phillip
; Gernann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/893,072
; FILING DATE: 15-Jul-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,749
; FILING DATE: 18-Aug-2000
; APPLICATION NUMBER: US/09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
```

```
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-893-072-25
Query Match 97.8%; Score 1725; DB 5; Length 379;
Best Local Similarity 88.1%; Pred. No. 1.6e-133;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;
QY 1 MAAAAQGGGGPRTEGVGPGVGEVEMVKGPFDVGPRTQLOYGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGVGPGVGEVEMVKGPFDVGPRTQLOYGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLRQIILLRFHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHQTYCQRTLRQIILLRFHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
DB 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHIL----- 258
DB 241 NRPIFGKHYLDQNLHILGLGSPSQEDLNCIINMKARNVLQSLPSKTKVAVAKLFPKSD 300
QY 259 --ALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLKPERL 316
DB 301 SKALDILLDRMLTFNPNKRITVEEALAHPLYEQYDPTDEPVAEEPTTFAMELDDLKPERL 360
QY 317 KELIFOETARFQGVLEAP 335
DB 361 KELIFOETARFQGVLEAP 379
RESULT 13
US-10-072-036-39
; Sequence 39, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BURON
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An Ir
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 39
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Erk1 fusion construct
US-10-072-036-39
Query Match 97.8%; Score 1725; DB 4; Length 631;
Best Local Similarity 88.1%; Pred. No. 3e-133;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;
QY 1 MAAAAQGGGGPRTEGVGPGVGEVEMVKGPFDVGPRTQLOYGEGAYGMVSSAY 60
```

Db 253 MAAAAAQQGGGPRRTGVPVGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAY 312
Qy 61 DHVKTRVAIKKISPEHQTYCORTLEIQILRFRHENVIGIRDIRASTLEAMRDVVI 120
Db 313 DHVKTRVAIKKISPEHQTYCORTLEIQILRFRHENVIGIRDIRASTLEAMRDVVI 372
Qy 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 373 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 432
Qy 181 KICDFGLARIADPEHDHGTFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
Db 433 KICDFGLARIADPEHDHGTFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 492
Qy 241 NRPIFGKHYLDQNLHIL----- 258
Db 493 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAAKLFPKSD 552
Qy 259 --ALDLDRLMTNPNKRIITVEBALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 316
Db 553 SKALDLDRLMTNPNKRIITVEBALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRL 612
Qy 317 KELIFQETARFQPGVLEAP 335
Db 613 KELIFQETARFQPGVLEAP 631

RESULT 14
US-10-072-036-57
; Sequence 57, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJRON
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; PRIOR FILING DATE: 2002-09-13
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mERK1-F64L-S65T-GFP fusion
US-10-072-036-57

Query Match 94.1%; Score 1659; DB 4; Length 624;
Best Local Similarity 85.4%; Pred. No. 8e-128;
Matches 323; Conservative 1; Mismatches 10; Indels 44; Gaps 1;

Qy 2 AAAAAQQGGGPRRTGVPVGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAYD 61
Db 3 AAAAAQQGGGPRRTGVPVGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAYD 62
Qy 62 HVKTRVAIKKISPEHQTYCORTLEIQILRFRHENVIGIRDIRASTLEAMRDVVI 121
Db 63 HVKTRVAIKKISPEHQTYCORTLEIQILRFRHENVIGIRDIRASTLEAMRDVVI 122
Qy 122 QDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK 181
Db 123 QDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLK 182
Qy 182 KICDFGLARIADPEHDHGTFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN 241
Db 183 KICDFGLARIADPEHDHGTFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLSN 242

Qy 242 RPIFGKHYLDQNLHIL----- 258
Db 243 RPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAAKLFPKSDS 302
Qy 259 -ALDLDRLMTNPNKRIITVEBALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRLK 317
Db 303 KALDLDRLMTNPNKRIITVEBALAHPLYEQYYDPTDEPVABEPTTFAMELDDLPKRLK 362
Qy 318 ELIFQETARFQPGVLEAP 335
Db 363 ELIFQETARFQGAPEGP 380

RESULT 15
US-10-461-402-8
; Sequence 8, Application US/10461402
; Publication No. US20030229209A1
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; Moller, Niels P.H.
; Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; KINASE, SEQUENCES, AND
; METHODS OF PRODUCTION
; AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/461,402
; FILING DATE: 16-Jun-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,953A
; FILING DATE: June 2, 1995
; APPLICATION NUMBER: 08/029,494
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 209/267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-461-402-8

Query Match 91.1%; Score 1606; DB 4; Length 355;
Best Local Similarity 87.6%; Pred. No. 9.4e-124;
Matches 311; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

Qy 25 PGEVEMVKQOPFDVGPRTYQLOVIGEGAYGMVSSAYDHVKTRVAIKKISPEHQTYCOR 84

Db	1	PGEVEMVKQPFDVGPRTQLOVIGBGAYGMVSSAYDHVRKTRVAIKKISPFHQTYCQR	60
Qy	85	TLREIQIILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC	144
Db	61	TLREIQIILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC	120
Qy	145	YFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLTEY	204
Db	121	YFLYQILRLGLKYTHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLTEY	180
Qy	205	VATRWYRAPEIMLNSKGYTKSIDINWSVGCILAEMLSNRPIFPKGHYLDQLNHIL-----	258
Db	181	VATRWYRAPEIMLNSKGYTKSIDINWSVGCILAEMLSNRPIFPKGHYLDQLNHILGILGSP	240
Qy	259	-----ALDLDRLMTFNPKNRITVEEA	280
Db	241	SOEDLNCIINMKARNYLQSLPSKTKVAMAKLFPKSDSKALDLDRLMTFNPKNRITVEEA	300
Qy	281	LAHPYLEQYYDPTDEPVASEPPTFAMELDDLPKERLKELI FORTARFQGVLEAP	335
Db	301	LAHPYLEQYYDPTDEPVASEPPTFAMELDDLPKERLKELI FORTARFQGVLEAP	355

Search completed: February 6, 2006, 15:40:56
Job time : 117.816 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:37:06 ; Search time 10.1369 Seconds
(without alignments)
387.254 Million cell updates/sec

Title: US-10-623-108-2

Perfect score: 1763

Sequence: 1 MAAAAAGGGGEPRETEGV.....LKELIFQETARFGVLEAP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pdb:*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pdb:*
3: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pdb:*
4: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pdb:*
5: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pdb:*
6: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pdb:*
7: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pdb:*
8: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1725	97.8	379	7	US-11-109-156-16
2	1444.5	81.9	360	6	US-10-878-556A-134
3	1444.5	81.9	360	7	US-11-186-284-115
4	747.5	42.4	360	6	US-10-886-329-1
5	678.5	38.5	365	6	US-10-770-726-69
6	620	35.2	422	6	US-10-857-780-21
7	620	35.2	422	7	US-11-127-817-18
8	620	35.2	426	7	US-11-127-817-20
9	620	35.2	464	7	US-11-127-817-19
10	464.5	26.3	294	7	US-11-024-959-263
11	464.5	26.3	302	7	US-11-024-959-262
12	461	26.1	304	7	US-11-024-959-264
13	460.5	26.1	294	7	US-11-024-959-265
14	454.5	25.8	460	7	US-11-024-959-384
15	450.5	25.6	520	7	US-11-024-959-272
16	449	25.5	277	7	US-11-127-817-21
17	447	25.4	298	6	US-10-770-726-51
18	442	25.1	795	6	US-10-770-726-49
19	434	24.6	292	6	US-10-770-726-53
20	432.5	24.5	297	6	US-10-770-726-48
21	432.5	24.5	297	7	US-11-109-156-11
22	432.5	24.5	346	6	US-10-770-726-55
23	427.5	24.2	463	7	US-11-024-959-395
24	418.5	23.7	509	7	US-11-024-959-393
25	414.5	23.5	555	7	US-11-024-959-270

RESULT 1
US-11-109-156-16
; Sequence 16, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-Ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-16

ALIGNMENTS

26	413.5	23.5	256	6	US-10-877-346-72	Sequence 72, Appl
27	413.5	23.5	256	7	US-11-113-424-183	Sequence 183, App
28	411	23.3	313	7	US-11-024-959-267	Sequence 267, App
29	410	23.3	496	6	US-10-770-726-72	Sequence 72, Appl
30	409	23.2	433	6	US-10-770-726-65	Sequence 65, Appl
31	408	23.1	476	7	US-11-024-959-385	Sequence 385, App
32	406	23.0	435	6	US-10-786-065-5	Sequence 5, Appli
33	402	22.8	469	7	US-11-024-959-271	Sequence 271, App
34	401	22.7	477	7	US-11-024-959-261	Sequence 261, App
35	401	22.7	483	6	US-10-451-375-12	Sequence 12, Appl
36	399	22.6	706	7	US-11-024-959-269	Sequence 269, App
37	397	22.5	330	6	US-10-786-065-8	Sequence 8, Appli
38	396	22.5	718	7	US-11-024-959-273	Sequence 273, App
39	395	22.4	330	6	US-10-786-065-10	Sequence 10, Appli
40	395	22.4	330	6	US-10-786-065-9	Sequence 9, Appli
41	394.5	22.4	583	7	US-11-024-959-268	Sequence 268, App
42	394	22.3	302	7	US-11-024-959-386	Sequence 386, App
43	389	22.1	845	7	US-11-024-959-391	Sequence 391, App
44	387.5	22.0	303	6	US-10-770-726-52	Sequence 52, Appl
45	381.5	21.6	305	7	US-11-024-959-266	Sequence 266, App

```
Query Match      97.8%; Score 1725; DB 7; Length 379;
Best Local Similarity 88.1%; Pred. No. 5.3e-147;
Matches 334; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAAAAAQQGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAQQGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||

QY 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||
Db 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTYYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 181 KICDFGLARIADPEHDHTGFLTYYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||

QY 241 NRPIPGKHLYDQNLNHL----- 258
   |||||
Db 241 NRPIPGKHLYDQNLNHL----- 258
   |||||

QY 259 --ALDILLDRMLTFNPNKRITVEALAHPLYEQYDPTDVPVABEPTTFAMELDDLPKERL 316
   |||||
Db 301 SKALDILLDRMLTFNPNKRITVEALAHPLYEQYDPTDVPVABEPTTFAMELDDLPKERL 360
   |||||

QY 317 KELIFOETARFQGVLEAP 335
   |||||
Db 361 KELIFOETARFQGVLEAP 379
   |||||

RESULT 2
US-10-878-556A-134
; Sequence 134, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/mk01_human
; DATABASE ENTRY DATE: 1992-12-01
US-10-878-556A-134

Query Match      81.9%; Score 1444.5; DB 6; Length 360;
Best Local Similarity 74.3%; Pred. No. 6.1e-122;
Matches 278; Conservative 18; Mismatches 17; Indels 61; Gaps 2;

QY 1 MAAAAAQQGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAAGAG-----PEMVRQVDFVGPRTYNTLSYIGEGAYGMVCSAY 43
   |||||

QY 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 44 DNVNKRVAIKKISPEHQTYCORTLREIKILLRFRHENIIGINDIIRAPTIEQMKDVYI 103
   |||||

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||
Db 104 VQDLMETDLYKLLKTQHLNSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 163
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTYYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 164 KICDFGLARVADPDHDHTGFLTYYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 223
   |||||
```

```
QY 241 NRPIPGKHLYDQNLNHL----- 258
   |||||
Db 224 NRPIPGKHLYDQNLNHLGILGPSQEDLNCIINLKARNYLLSLPHKNKVPWNLFPNAD 283
   |||||

QY 259 --ALDILLDRMLTFNPNKRITVEALAHPLYEQYDPTDVPVABEPTTFAMELDDLPKERL 316
   |||||
Db 284 SKALDILLDRMLTFNPNKRITVEALAHPLYEQYDPSDEPTAEAFKFDMLDDLPKEXL 343
   |||||

QY 317 KELIFOETARFQGP 330
   |||||
Db 344 KELIFEETARFQGP 357
   |||||

RESULT 3
US-11-186-284-115
; Sequence 115, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-115

Query Match      81.9%; Score 1444.5; DB 7; Length 360;
Best Local Similarity 74.3%; Pred. No. 6.1e-122;
Matches 278; Conservative 18; Mismatches 17; Indels 61; Gaps 2;

QY 1 MAAAAAQQGGGGPRRTGEGVGPVGEVEMVKGQPFDPVGPRTYQLOYIGEGAYGMVSSAY 60
   |||||
Db 1 MAAAAAAGAG-----PEMVRQVDFVGPRTYNTLSYIGEGAYGMVCSAY 43
   |||||

QY 61 DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||||
Db 44 DNVNKRVAIKKISPEHQTYCORTLREIKILLRFRHENIIGINDIIRAPTIEQMKDVYI 103
   |||||

QY 121 VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
   |||||
Db 104 VQDLMETDLYKLLKTQHLNSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 163
   |||||

QY 181 KICDFGLARIADPEHDHTGFLTYYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 240
   |||||
Db 164 KICDFGLARVADPDHDHTGFLTYYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS 223
   |||||

QY 241 NRPIPGKHLYDQNLNHL----- 258
   |||||
Db 224 NRPIPGKHLYDQNLNHLGILGPSQEDLNCIINLKARNYLLSLPHKNKVPWNLFPNAD 283
   |||||
```


QY 259 --ALDLDRLMTNPNKRTVEBALAHPLYEQYDPTDPEVAREPPTFAMELDDLPKRL 316
Db 284 SKALDLDRLMTNPNKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFOMELDLPKRL 343
QY 317 KELIFOETARFQPG 330
Db 344 KELIFETARFQPG 357

RESULT 4

US-10-886-329-1
; Sequence 1, Application US/10886329
; Publication No. US20050288286A1
; GENERAL INFORMATION:
; APPLICANT: Deciphra Pharmaceuticals, Inc.
; APPLICANT: Flynn, Daniel L
; APPLICANT: Petillo, Peter A
; TITLE OF INVENTION: Anti-Inflammatory Medicaments
; FILE REFERENCE: 34477-CIP
; CURRENT APPLICATION NUMBER: US/10/886,329
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: US 60/437,415
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US 10/746,460
; PRIOR FILING DATE: 2003-12-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-886-329-1

Query Match 42.4%; Score 747.5; DB 6; Length 360;
Best Local Similarity 45.1%; Pred. No. 1.3e-59;
Matches 155; Conservative 55; Mismatches 81; Indels 53; Gaps 6;
QY 36 FDVGPRTQIQYIGEGAYGMVSSAYDHVRKTRVAIKKIS-PFEHQTYCORTLREIQILLR 94
Db 18 WEVPERYQNLSPVSGAYGVSACAFTKTLRVAVKLSRPFQSIHAKRTYRELRLKH 77
QY 95 FRHENVIGIRDLR-ASTLEAMRDVYVQDLMETDLYKLLKSQOLSNDHICYFLYQILRG 153
Db 78 MKHENVIGLDDVETPARSLSEBDFNVDVYLVTLMGADLNNIVKQKLTDDHVFQILYQILRG 137
QY 154 LKVIHSANVLRDLKPSNLINTTCDLKICDFGLARIADPEHDTGFLTEYVATRWYRAP 213
Db 138 LKVIHSADIIHRDLKPSNLAVNEDCELKILDFGLARHTDDE-----MTGYVATRWYRAP 191
QY 214 EIMLSNGYTKSIDIWSVGCILAEMLSNRPIFGKHLYDQ----- 254
Db 192 EIMLNNHYNQTVDIWSVGCIMAEMLTGTTLFGTDDHIDQLKILRLVGTGPAELLKKIS 251
QY 255 -----NHI-----LALDLDRLMTNPNKRTVEBALAHPLYEQY 289
Db 252 SESARNYIQSLTKPMKPNFANVFICANPLAVDLLEKMLVDSKRTITAAQALAHAYPAQY 311
QY 290 YDPTDEVAEPPTFAMELDDLPKRLKELIFOETARFQPGVLS 333
Db 312 HDPDPEVA-DPYDQSPESRDLIDEWKSLTYDEVISFVPPPLD 354

RESULT 5

US-10-770-726-69
; Sequence 69, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM101079 (031896-010000)

; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 69
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-69

Query Match 38.5%; Score 678.5; DB 6; Length 365;
Best Local Similarity 41.6%; Pred. No. 2e-53;
Matches 139; Conservative 55; Mismatches 87; Indels 53; Gaps 5;
QY 42 YTQLOVIGEGAYGMVSSAYDHVRKTRVAIKKIS-PFEHQTYCORTLREIQILLRPFHENV 100
Db 25 YVSPTHVSGAYGVSACAFTKTLRVAVKLSRPFQSEIFAKRAYRELLLLKHMQHENV 84
QY 101 IGIRDLR-ASTLEAMRDVYVQDLMETDLYKLLKSQOLSNDHICYFLYQILRGKLYIHS 159
Db 85 IGLDVTTPASSLRNPFYLYVNPFPMTDQKIM-GMEFSEBEKIQLVYVQMLKGLYIHS 143
QY 160 ANVLRDLKPSNLINTTCDLKICDFGLARIADPEHDTGFLTEYVATRWYRAPPEIMLS 219
Db 144 AGVHRDLKPSNLAVNEDCELKILDFGLARHADA-----MTGYVATRWYRAPPEVILSW 197
QY 220 KGYTKSIDIWSVGCILAEMLSNRPIFGKHLYDQ----- 258
Db 198 MHNQTVDIWSVGCIMAEMLTGTTLFGKDYLDQTLTKLKVTPGPGTEFVQKLNDAKAS 257
QY 259 -----ALDLDRLMTNPNKRTVEBALAHPLYEQYDPTDE 295
Db 258 YTQSLPQTPRKDFTQLFPRASQAADLEKMLEVDVKRLTAAQALTHPFFEFPRDPEE 317
QY 296 PVAEBEPTFAMELDDLPKRLKELIFOETARFQ 329
Db 318 TEAQPFDDSLSEKLTVDENKQHIYEIVNFSF 351

RESULT 6

US-10-857-780-21
; Sequence 21, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-780-21

Query Match 35.2%; Score 620; DB 6; Length 422;
Best Local Similarity 38.2%; Pred. No. 4.2e-48;
Matches 133; Conservative 67; Mismatches 80; Indels 68; Gaps 9;

QY 36 FDVGPRTQOYIGEGAYGMVSSAYDHVRKTRVAIKKIS--PFEHOTYQORTLREIQILLR 94
Db 58 FTVLKRYQNLKPTIGSGAQGIVCAAYDAVLDNRVAIKLSRPFQNTQAKRAYRELVLKMC 117
QY 95 FHENVIGIRDILR--ASTLEAMRDVYIVODLMETDLYKLLKSQQLSNDHICYFLYQILRG 153
Db 118 VNHKNIISLLNVFTPKTLEEFQDVLVLMELMDANLCQVIQ--MELDHMSYLLYQMLCG 176
QY 154 LKYIHSANVLHRLDKPSNLLINTTCDLKI CDGLARIADPEHDHTGF--LTEYVATRWYRA 212
Db 177 IKHLHSAGIIHRLDKPSNIVKSDCTLKILDFGLARTAG-----TSFMWTPYVVTYYRA 231
QY 213 PETMNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHVLDQNLNHL----- 258
Db 232 PEVIL--GMGYKENVDIWSVGCINGEMVRHKILFPGRDYIDQWKNVIEQGTGCPPEFMKKL 290
QY 259 -----ALDLDRLMLTFNPNKRITV 277
Db 291 OPTVRNVENRPKYAGLTFFPKLFPDSLFPADSEHNKLKASQARDLLSKMLVIDPAKRISV 350
QY 278 BEALAHPLYEQYDPTDPEVAEPTTFAMELDDLKP--ERLKELIFQE 323
Db 351 DDALQHPYINWYDPA--EVEAPPPQIYDKQDREHTEIEWKELIYKE 397

RESULT 7

US-11-127-817-18
; Sequence 18, Application US/11127817
; Publication No. US20050287519A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann, Pascal G.
; APPLICANT: Spittaels, Marcel
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions And Compound Assays For Inhibiting
; FILE OF INVENTION: Amyloid-Beta Protein Production
; CURRENT APPLICATION NUMBER: US/11/127,817
; PRIOR FILING DATE: 2005-05-12
; PRIOR FILING DATE: 2004-05-12
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 534
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 18
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-817-18

Query Match 35.2%; Score 620; DB 7; Length 422;
Best Local Similarity 38.2%; Pred. No. 4.2e-48;
Matches 133; Conservative 67; Mismatches 80; Indels 68; Gaps 9;

QY 36 FDVGPRTQOYIGEGAYGMVSSAYDHVRKTRVAIKKIS--PFEHOTYQORTLREIQILLR 94
Db 58 FTVLKRYQNLKPTIGSGAQGIVCAAYDAVLDNRVAIKLSRPFQNTQAKRAYRELVLKMC 117
QY 95 FHENVIGIRDILR--ASTLEAMRDVYIVODLMETDLYKLLKSQQLSNDHICYFLYQILRG 153
Db 118 VNHKNIISLLNVFTPKTLEEFQDVLVLMELMDANLCQVIQ--MELDHMSYLLYQMLCG 176
QY 154 LKYIHSANVLHRLDKPSNLLINTTCDLKI CDGLARIADPEHDHTGF--LTEYVATRWYRA 212
Db 177 IKHLHSAGIIHRLDKPSNIVKSDCTLKILDFGLARTAG-----TSFMWTPYVVTYYRA 231
QY 213 PETMNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHVLDQNLNHL----- 258
Db 232 PEVIL--GMGYKENVDIWSVGCINGEMVRHKILFPGRDYIDQWKNVIEQGTGCPPEFMKKL 290
QY 259 -----ALDLDRLMLTFNPNKRITV 277

Db 291 OPTVRNVENRPKYAGLTFFPKLFPDSLFPADSEHNKLKASQARDLLSKMLVIDPAKRISV 350
QY 278 BEALAHPLYEQYDPTDPEVAEPTTFAMELDDLKP--ERLKELIFQE 323
Db 351 DDALQHPYINWYDPA--EVEAPPPQIYDKQDREHTEIEWKELIYKE 397

RESULT 8

US-11-127-817-20
; Sequence 20, Application US/11127817
; Publication No. US20050287519A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann, Pascal G.
; APPLICANT: Spittaels, Marcel
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions And Compound Assays For Inhibiting
; FILE OF INVENTION: Amyloid-Beta Protein Production
; CURRENT APPLICATION NUMBER: US/11/127,817
; PRIOR FILING DATE: 2005-05-12
; PRIOR FILING DATE: 2004-05-12
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 534
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 20
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-817-20

Query Match 35.2%; Score 620; DB 7; Length 426;
Best Local Similarity 38.2%; Pred. No. 4.3e-48;
Matches 133; Conservative 67; Mismatches 80; Indels 68; Gaps 9;

QY 36 FDVGPRTQOYIGEGAYGMVSSAYDHVRKTRVAIKKIS--PFEHOTYQORTLREIQILLR 94
Db 20 FTVLKRYQNLKPTIGSGAQGIVCAAYDAVLDNRVAIKLSRPFQNTQAKRAYRELVLKMC 79
QY 95 FHENVIGIRDILR--ASTLEAMRDVYIVODLMETDLYKLLKSQQLSNDHICYFLYQILRG 153
Db 80 VNHKNIISLLNVFTPKTLEEFQDVLVLMELMDANLCQVIQ--MELDHMSYLLYQMLCG 138
QY 154 LKYIHSANVLHRLDKPSNLLINTTCDLKI CDGLARIADPEHDHTGF--LTEYVATRWYRA 212
Db 139 IKHLHSAGIIHRLDKPSNIVKSDCTLKILDFGLARTAG-----TSFMWTPYVVTYYRA 193
QY 213 PETMNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHVLDQNLNHL----- 258
Db 194 PEVIL--GMGYKENVDIWSVGCINGEMVRHKILFPGRDYIDQWKNVIEQGTGCPPEFMKKL 252
QY 259 -----ALDLDRLMLTFNPNKRITV 277
Db 253 OPTVRNVENRPKYAGLTFFPKLFPDSLFPADSEHNKLKASQARDLLSKMLVIDPAKRISV 312
QY 278 BEALAHPLYEQYDPTDPEVAEPTTFAMELDDLKP--ERLKELIFQE 323
Db 313 DDALQHPYINWYDPA--EVEAPPPQIYDKQDREHTEIEWKELIYKE 359

RESULT 9

US-11-127-817-19
; Sequence 19, Application US/11127817
; Publication No. US20050287519A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann, Pascal G.
; APPLICANT: Spittaels, Marcel
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions And Compound Assays For Inhibiting
; FILE OF INVENTION: Amyloid-Beta Protein Production

```

; FILE REFERENCE: P27,800-D USA
; CURRENT APPLICATION NUMBER: US/11/127,817
; CURRENT FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 534
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 19
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-127-817-19

Query Match 35.2%; Score 620; DB 7; Length 464;
Best Local Similarity 38.2%; Pred. No. 4.8e-48;
Matches 133; Conservative 67; Mismatches 80; Indels 68; Gaps 9;

QY 36 FDVGPRTQYIGEGAYGMVSSAYDHVRKTRVAIKKIS-PFEHQTYCQRTLRRIQILRRHEN 94
Db 58 FTVLKRYQNLKPIGSAQGVCAAYDAVLDNRNVAIKKLSRPFQNTAKRAYRELVMKC 117

QY 95 FPHNVIGIRDILR-ASTLEAMRDVIVQDLMETDLYKLLKSQQLSNDHICVLYQILRG 153
Db 118 VNHKNIISLNVPTPQKTLFEQDQVYLVMEIDANLCQVIQ-MELDHERRMSYLLYQMLCG 176

QY 154 LKVIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGF-LTEYVATRWYRA 212
Db 177 IKHLISAGIIRDLKPSNLLINTTCDLKICDFGLARTAG-----TSFMTPTVYVTRYRA 231

QY 213 PEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYLDQLNHIL----- 258
Db 232 PEVIL-GMGYKENVDIWSVGCINGEMVRHKILPGRDIYDQWKNVIEQLGTPCPPEFMKCL 290

QY 259 -----ALDLDRLMTFNPKNRITVEEALAHAYLE 287
Db 291 QPTVRNVENRKYAGLTPKLPDLSLPADSEHNKLKASQARDLSKMLVIDPAKRISV 350

QY 278 EEALAHAYLQYDPTDEPAEPPFTFAMELDPK--ERLKELIFQE 323
Db 351 DDALQHPYINWYDPA-EVEAPPQIYDKQDEREHTIEWKSLIYKE 397

RESULT 10
US-11-024-959-263
; Sequence 263, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 04463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 263
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
; US-11-024-959-262

Query Match 26.3%; Score 464.5; DB 7; Length 302;
Best Local Similarity 35.6%; Pred. No. 2.2e-34;
Matches 105; Conservative 46; Mismatches 87; Indels 57; Gaps 6;

QY 41 RYTQLOQYIGEGAYGMVSSAYDHVRKTRVAIKKIS-PFEHQTYCQRTLRRIQILRRHEN 99
Db 3 KYQQLAKIGEGTYGIVYKADKKGSGELLALKIRLEADEGIPSTAIRSISLKLQHPN 62

QY 100 VIGIRDILRASTLEAMRDVIVQDLMETDLYKLLKS---QQLSNDHICVLYQILRGKY 156
Db 63 IVRLYDVVHTE-----KGLTVFEFLDQDLKKYLDACDNGLEPYTVKSLYQLLQGIAP 117

QY 157 IHSANVLRDLKPSNLLINTTCDLKICDFGLAR-IADPEHDHTGFTEYVATRWYRAPEI 215
Db 118 CHEHRLVLRDLKPSNLLINWEGELKLADFGARAFGIPVRNY---THEVVTWYRAPDV 173

QY 216 MLNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYLDQLNHIL----- 258
Db 174 LMGRKYSTQVDIWSVGCIFAEMVNGRPLPFGSGEQDLRIKFTLGTPLSKTWPGMAEL 233

QY 259 -----ALDLDRLMTFNPKNRITVEEALAHAYLE 287
Db 234 PDFKDNFPKYVQSFKKICPKKLDLGLDLSRLMLQYDPAKRISASQAMGHYPFK 288

Matches 108; Conservative 45; Mismatches 80; Indels 67; Gaps 8;

QY 41 RYTQLOQYIGEGAYGMVSSAYDHVRKTRVAIKKIS-PFEHQTYCQRTLRRIQILRRHEN 99
Db 3 QYEKIEKIGEGTYGVVYKAIDRSTNKTIALKIRLEQDEGVPSTAIRSISLKLQHPN 62

QY 100 VIGIRDILRASTLEAMRDVIVQDLMETDLYKLLKS-QQLSND--HICVLYQILRGKY 156
Db 63 IVKLDQVHSE-----RRLYLVFEFLDQDLKKHMSQCPFSKDTHTIKMFLYQLLKGISY 117

QY 157 IHSANVLRDLKPSNLLIN-TTCDLKICDFGLARIADPEHDHTGF-----LTEYVATRW 209
Db 118 CHSHRVLRDLKPSNLLIDRRTNSLKLADFGAR-----AFGIPVTRTPTHEVVTW 168

QY 210 YRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYLDQLNHIL----- 258
Db 169 YRAPEILLGSRHYSTPVDVWSVGCIFAEMVNRPLPFGDSEIDELFKIFRIMGTPNEDSW 228

QY 259 -----ALDLDRLMTFNPKNRITVEEALAHAYLE 287
Db 229 PGVTSLPDPFKSTFPKWASQDLKTVTPTVDPAGIDLLSKMLCWDPRRITAKVALEHEYFK 288

RESULT 11
US-11-024-959-262
; Sequence 262, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 04463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 262
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
; US-11-024-959-262

Query Match 26.3%; Score 464.5; DB 7; Length 302;
Best Local Similarity 35.6%; Pred. No. 2.2e-34;
Matches 105; Conservative 46; Mismatches 87; Indels 57; Gaps 6;

QY 41 RYTQLOQYIGEGAYGMVSSAYDHVRKTRVAIKKIS-PFEHQTYCQRTLRRIQILRRHEN 99
Db 3 KYQQLAKIGEGTYGIVYKADKKGSGELLALKIRLEADEGIPSTAIRSISLKLQHPN 62

QY 100 VIGIRDILRASTLEAMRDVIVQDLMETDLYKLLKS---QQLSNDHICVLYQILRGKY 156
Db 63 IVRLYDVVHTE-----KGLTVFEFLDQDLKKYLDACDNGLEPYTVKSLYQLLQGIAP 117

QY 157 IHSANVLRDLKPSNLLINTTCDLKICDFGLAR-IADPEHDHTGFTEYVATRWYRAPEI 215
Db 118 CHEHRLVLRDLKPSNLLINWEGELKLADFGARAFGIPVRNY---THEVVTWYRAPDV 173

QY 216 MLNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYLDQLNHIL----- 258
Db 174 LMGRKYSTQVDIWSVGCIFAEMVNGRPLPFGSGEQDLRIKFTLGTPLSKTWPGMAEL 233

QY 259 -----ALDLDRLMTFNPKNRITVEEALAHAYLE 287
Db 234 PDFKDNFPKYVQSFKKICPKKLDLGLDLSRLMLQYDPAKRISASQAMGHYPFK 288

```

```
RESULT 12
US-11-024-959-264
; Sequence 264, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 264
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-264

Query Match      26.1%; Score 461; DB 7; Length 304;
Best Local Similarity 35.2%; Pred. No. 4.6e-34;
Matches 105; Conservative 46; Mismatches 91; Indels 56; Gaps 6;

QY 41 RYTQLOYIGEGAYGMVSSAYDHVTRVVAIKKIS-PFEHQTYCQRTLRREIQILLRPHEN 99
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 KYEKLKEGTYGVVYKAQDKTKTEIYALKKIRLESEDEGIPSTAIRAIALLKELQHPN 68
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 100 VTGIRDILRASTLEAMRDVYVODLMETDLYKLLKKSQQLSNDH--ICVFLYQILRLGKYI 157
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 VRIHDVIHTN-----KGLILFEFVDYDLKFLHFDGIDPKIVKSLYQLVRGVAHC 123
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 158 HSANVLHRLDKPSNLLINTTCDLKICDFGLAR--IADPEHDHTGFLFEYVATRWYRAPEIM 216
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 HQQKVLHRLDKPQNLLVSQEGILKLGDFGLARAFGIPVKNY---TNEVVTLYWRAPDIL 179
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 217 LNSKGYTKSIDWSVGCILAEMLSNRPIPPGKHVLDQLNHL----- 258
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 LGSKNYSTSVDIWSIGCIFVEMLNKPLFPGSSEQDQLKKIKFIMGTPDATKWPGLAELP 239
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 259 -----ALDILDRMLTFNPNKRITVEEALAHPLYEQYD 291
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 DWKPNFEKYPGEPLNKKVCPKMDPDGLDLDLKKCNPSERIAAKNMGSHPTFKDIPD 297
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-11-024-959-265
; Sequence 265, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
```

```
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 265
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-265

Query Match      26.1%; Score 460.5; DB 7; Length 294;
Best Local Similarity 35.7%; Pred. No. 4.8e-34;
Matches 107; Conservative 45; Mismatches 81; Indels 67; Gaps 8;

QY 41 RYTQLOYIGEGAYGMVSSAYDHVTRVVAIKKIS-PFEHQTYCQRTLRREIQILLRPHEN 99
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 QYEKVEKIGEGTYGVVYKAIDRLTNETIALLKIRLEQDEGVSPSTAIRAISLLKEMQHGN 62
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 100 VIGIRDILRASTLEAMRDVYVODLMETDLYKLLKKSQQLSNDH--ICVFLYQILRLGKY 156
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 IVRLQDVVHSE-----NRLYLVEFYLDLCKHMDSPDFAKDPRLVKIFLYQILRGYAT 117
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 157 IHSANVLHRLDKPSNLLIN-TTCDLKICDFGLARIADPEHDHTGF-----LTEYVATRW 209
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 CHSHRVLHRLDKPQNLLIDRRTNALKLADFGLAR-----AFGIPVTFTHEVVTLM 168
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 210 YRAPEIMLNSKGYTKSIDWSVGCILAEMLSNRPIPPGKHVLDQLNHL----- 258
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 YRAPEILLGSRHYSTPVDVWSVGCIFAEWVNQRPFGDSEIDELFKIFRILGTPTNEDTW 228
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 259 -----ALDILDRMLTFNPNKRITVEEALAHPLYE 287
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 PGVTALPDFKSAPKWPAPKNLQDMVPGVGLNSAGIDLLSKMLCLDPSKRITARSALHEVFK 288
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
US-11-024-959-384
; Sequence 384, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 384
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-384

Query Match      25.8%; Score 454.5; DB 7; Length 460;
Best Local Similarity 35.1%; Pred. No. 3e-33;
Matches 104; Conservative 53; Mismatches 72; Indels 67; Gaps 9;

QY 41 RYTQLOYIGEGAYGMVSSAYDHVTRVVAIKKISPEHQTYCQR---TLREIQILLRPH 97
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 RYKVKELGDGTGYGVWVKALNQOOTHETHEIVAIKQK---KRYIWEBCINREKSLRLNH 59
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 98 ENVIGIRDILRASTLEAMRDVYVODLMETDLYKLLK--SQQLSNDHICYFLYQILRLGK 155
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 PNIILKKEVIREN-----NELEFFIYMECNLYQIMKGRSTPSETAIKFCVQILQGLS 114
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 156 YIHSANVLHRLDKPSNLLINTTCDL-KICDFGLAR--IADPEHDHTGFLTEYVATRWYRA 212
```

```
Db 115 YMRNGYFHRDLKPENLLV--TSLIKIADFGlareVLTSPY-----TDYVSTRWYRA 166
Qy 213 PEIMNSKGYTKSIDISWVCILAEMLSNRPIPPGKHLDQLNHIL-----258
Db 167 PEVLQSPYTTAIDMWAVGAILAELFTLHPLPGESELDIYKICGVLGTPDYETWPDG 226
Qy 259 -----ALDLLDRMLTFNPNKRITVEEALAHPY 285
Db 227 MQLAAFRNFIFPQPLPVNLSVLIPHASPEAIDLITRLCSWDPOKRPATAEQALHPF 282
```

RESULT 15

```
US-11-024-959-272
; Sequence 272, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 272
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-272
```

```
Query Match 25.6%; Score 450.5; DB 7; Length 520;
Best Local Similarity 33.6%; Pred. No. 8e-33;
Matches 116; Conservative 53; Mismatches 105; Indels 75; Gaps 8;

Qy 42 YTLQYIGEGAYGMVSSAYDHRKTRVAIKI-SPFEHQTYCQRTLEIOILLRFRHNV 100
Db 25 FEKLEQIGEGTYGVYNAKKTGEIVALKKIRMDNEREGFPITAREIKILKKLHENV 84

Qy 101 IGIRDIILRASTLEMRD-----VIIVQDLMETDLYKLL--KSQQLSNDHICYFL 147
Db 85 IKLKEIVTSPGPEKDEQGRPEGNKYKGGIYVFEYMDHDLTGLADRFPMRFPVQIKCYM 144

Qy 148 YQILRGLKYTHSANVLRDLKPSNLLINTTCDLKIADFGlareVLTSPYVAT 207
Db 145 RQLITGLHYCHINQVLRDIDKGNLLIDNEGKLDADFGAR--SFSNDENANLTNRVIT 202

Qy 208 RWRAPETIMNSKGYTKSIDISWVCILAEMLSNRPIPPGKHLDQLNHIL-----258
Db 203 LWYRPPELLGATKYGPAVDWMSVGCIFAEILHGKPIFFPGKDEPEQLNKIFELCGAPDEI 262

Qy 259 -----ALDLLDRMLTFNPNKRITVEEALAHPY 284
Db 263 NWPGVSKIPWYNFKNRRLRLEVPFRHFRDRHLELLERMLTDFPSQISAKDALD-- 320

Qy 285 YLEQYV-----DPTDEPVAEEPTFAMELDDLPKELKELIFQETARFQ 328
Db 321 --AEYFWADPLPCDPKSLPKYESSHEFQTK-----KKRQQORQHEETAKRQ 364
```

Search completed: February 6, 2006, 15:41:22
Job time : 11.1369 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:19:31 ; Search time 286.062 Seconds
(without alignments)
551.408 Million cell updates/sec

Title: US-10-623-108-4
Perfect score: 1879
Sequence: 1 MAAAAAGGGGGPRRTG.....LKELIFQTARFQGVLEAP 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	100.0	359	9	ADW12903 Human SMA
2	1879	100.0	359	9	ADW12907 Human SMA
3	1873	99.7	359	6	ABU54596 Human NOV
4	1859	98.9	379	2	AAM15506 Mitogen a
5	1859	98.9	379	4	AAM38754 Human pol
6	1859	98.9	379	6	ABR43898 Human ERK
7	1859	98.9	379	7	ADP45050 Human kin
8	1859	98.9	379	7	ADH59632 Erk1 prot
9	1859	98.9	379	8	ADO22522 Biochemic
10	1859	98.9	379	8	ADQ37846 Human pho
11	1859	98.9	379	9	ADV42070 Human ext
12	1859	98.9	379	9	ADW12905 Human SMA
13	1859	98.9	379	9	ADZ65036 Mitogen a
14	1859	98.9	403	4	AM40540 Human pol
15	1853	98.6	379	2	AA142413 Extracell
16	1853	98.6	379	4	AAAG67618 Amino aci
17	1853	98.6	379	4	AAAG67439 Amino aci
18	1853	98.6	379	5	ABG31847 Human ext
19	1853	98.6	379	6	ABU89742 protein d
20	1853	98.6	379	7	ABR62625 Human pro
21	1853	98.6	379	8	ADO24425 Human PRO
22	1853	98.6	379	9	ADZ10062 Human bre
23	1853	98.6	631	2	AAW85006 Erk1-gree
24	1853	98.6	631	3	AA170778 EGFP-Erk1

25	1787	95.1	624	2	AAW85015	Aaw85015	Erk1-gree
26	1777	94.6	375	6	ABP96053	Abp96053	Human pro
27	1735	92.3	356	2	AAW48013	Aam48013	Extracell
28	1734	92.3	355	7	ADC64458	Adc64458	Human ext
29	1734	92.3	355	8	ADG98194	Adg98194	Human ext
30	1723	91.7	365	2	AA820103	Aar20103	ERK1. 8/2
31	1614.5	85.9	340	8	ABM82549	Abm82549	Human dia
32	1614.5	85.9	340	8	ABM82547	Abm82547	Human dia
33	1599	85.1	335	8	ADH48364	Adh48364	Human KPP
34	1599	85.1	335	8	ABM82550	Abm82550	Human dia
35	1599	85.1	335	8	ABM82548	Abm82548	Human dia
36	1599	85.1	335	9	ADW12901	Adw12901	Human SMA
37	1548.5	82.4	380	7	ADC64459	Adc64459	Human ext
38	1548.5	82.4	380	7	ADG98195	Adg98195	Human ext
39	1538.5	81.9	360	5	ABG31848	Abg31848	Human cer
40	1538.5	81.9	360	6	ABR92109	AbR92109	Human cer
41	1538.5	81.9	360	6	ABR43899	AbR43899	Human ERK
42	1538.5	81.9	360	7	ADP45051	Adh45051	Human kin
43	1538.5	81.9	360	7	ADH59634	Adh59634	Erk2 prot
44	1538.5	81.9	360	8	ADJ66608	Adj66608	ERK prote
45	1538.5	81.9	360	8	ADP56268	Adp56268	Human PRO

ALIGNMENTS

RESULT 1
ADW12903
ID ADW12903 standard; protein; 359 AA.
XX
AC ADW12903;
XX
DT 07-APR-2005 (first entry)
XX
DE Human SMAPK3V2 variant protein.
XX
KW DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW enzyme.
XX
XX Homo sapiens.
XX
PN US2005013817-A1.
XX
PD 20-JAN-2005.
XX
PF 18-JUL-2003; 2003US-00623108.
XX
PR 18-JUL-2003; 2003US-00623108.
XX
PA (DAIK/) DAI K.
XX
PI Dai K;
XX
DR WPI; 2005-080923/09.
DR N-PSDB; ADW12902.
XX
PT New isolated SMAPK3 polypeptides and nucleic acids, useful for diagnosing
PT diseases, e.g. cancers, associated with the deficiency of the SMAPK3 gene
PT in a mammal.
XX
PS Claim 1; SEQ ID NO 4; 55pp; English.
XX
CC The invention relates to novel isolated polypeptide (I) comprising an
CC amino acid sequence selected from sequences comprising 335 or 359 amino
CC acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
CC polynucleotides are useful for diagnosing diseases, e.g. cancers,
CC associated with the deficiency of the SMAPK3 gene in a mammal. The
CC fragments of the polypeptides and polynucleotides can also be used as
CC primers or probes. This sequence corresponds to the SMAPK3V1 variant
XX protein.
SQ Sequence 359 AA;

Query Match	100.0%;	Score 1879;	DB 9;	Length 359;
Best Local Similarity	100.0%;	Pred. No. 8.2e-187;		
Matches 359;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	MAAAAAGGGGGPRRTGEGVGPVGEVMVGQPFDDGPRYTQIQYIGEGAYGMWSAY	60		
Db 1	MAAAAAGGGGGPRRTGEGVGPVGEVMVGQPFDDGPRYTQIQYIGEGAYGMWSAY	60		
Qy 61	DHVRKTRVAIKKISPFESHQYTCQRTLREIQILLFRRHENVIGIRDILRASTLEAMRDVYI	120		
Db 61	DHVRKTRVAIKKISPFESHQYTCQRTLREIQILLFRRHENVIGIRDILRASTLEAMRDVYI	120		
Qy 121	VQDLMETDLYKLKSKQQLSNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL	180		
Db 121	VQDLMETDLYKLKSKQQLSNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDL	180		
Qy 181	KICDFGLIARADPEHDDHTGFLTEVATRWYRAPEIMLNSKGYSKSIDIWSVGCTIAEMLS	240		
Db 181	KICDFGLIARADPEHDDHTGFLTEVATRWYRAPEIMLNSKGYSKSIDIWSVGCTIAEMLS	240		
Qy 241	NRPIFPCKHYLDQLNHLIGILGSPSQEDLNCCIINNKAARYLQSLPSKTKYAWAKLFPKSD	300		
Db 241	NRPIFPCKHYLDQLNHLIGILGSPSQEDLNCCIINNKAARYLQSLPSKTKYAWAKLFPKSD	300		
Qy 301	SKALDLDLRMLTFNPNKRIITVAESPFTFAMELDDLPKRLKELIFQETARPQGVLEAP	359		
Db 301	SKALDLDLRMLTFNPNKRIITVAESPFTFAMELDDLPKRLKELIFQETARPQGVLEAP	359		

RESULT 2	
ADW12907	
ID	ADW12907 standard; protein; 359 AA.
XX	
AC	ADW12907;
XX	
DT	07-APR-2005 (first entry)
XX	
DE	Human SNAPK3V4 variant protein.
XX	
XX	DNA purification; diagnosis; cancer; mitogen-activated protein kinase;
KW	enzyme.
KW	
XX	
OS	Homo sapiens.
XX	
PN	US2005013817-A1.
XX	
PD	20-JAN-2005.
XX	
XX	18-JUL-2003; 2003US-00623108.
PF	
XX	
PR	18-JUL-2003; 2003US-00623108.
XX	
PA	(DAIK/) DAI K.
XX	
PI	Dai K;
XX	
DR	WPI; 2005-080923/09.
DR	N-PSDB; ADW12906.
XX	
PT	New isolated SNAPK3 polypeptides and nucleic acids, useful for diagnosing
PT	diseases, e.g. cancers, associated with the deficiency of the SNAPK3 gene
PT	in a mammal.
XX	
PS	Disclosure; SEQ ID NO 8; 55pp; English.
XX	
CC	The invention relates to novel isolated polypeptide (I) comprising an
CC	amino acid sequence selected from sequences comprising 335 or 359 amino
CC	acids (ADW12901 or ADW12903) or its fragments. The polypeptides and
CC	polynucleotides are useful for diagnosing diseases, e.g. cancers,
CC	associated with the deficiency of the SNAPK3 gene in a mammal. The
CC	fragments of the polypeptides and polynucleotides can also be used as
CC	primers or probes. This sequence corresponds to the SNAPK3V1 variant
CC	protein.

XX	SQ	Sequence 359 AA;	Query Match	100.0%;	Score 1879;	DB 9;	Length 359;
			Best Local Similarity	100.0%;	Pred. No. 8.2e-187;		
			Matches 359;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAAAAAGGGGGGPRTEGVGPGEVEMVKQPPFDVGPRTYQLOVIGGAYGMVSSAY	60				
Db	1	MAAAAAGGGGGGPRTEGVGPGEVEMVKQPPFDVGPRTYQLOVIGGAYGMVSSAY	60				
Qy	61	DHVKTRVAIKTISPEHQTYCORTLREIOILLRFRHENVIGIRDILLRASTLEAMRDVYI	120				
Db	61	DHVKTRVAIKTISPEHQTYCORTLREIOILLRFRHENVIGIRDILLRASTLEAMRDVYI	120				
Qy	121	VQDLMETDLYKLKSOQLSNDHI CYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL	180				
Db	121	VQDLMETDLYKLKSOQLSNDHI CYFLYQILRGLKYIHSANVLHRDLKPSNLLINTTCDL	180				
Qy	181	KICDQGLIARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS	240				
Db	181	KICDQGLIARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDIWSVGCILAEMLS	240				
Qy	241	NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSD	300				
Db	241	NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSD	300				
Qy	301	SKALDILLDRMLTFNPKNRIITVAEPPFTFAMELDDLPKERLKELIFOETARFQGVLEAP	359				
Db	301	SKALDILLDRMLTFNPKNRIITVAEPPFTFAMELDDLPKERLKELIFOETARFQGVLEAP	359				
RESULT 3							
ABU54596							
ID	ABU54596	standard; protein; 359 AA.					
XX	AC	ABU54596;					
XX	DT	03-JUN-2003 (first entry)					
XX	DE	Human NOVX polypeptide #55.					
XX	KW	Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;					
XX	KW	hypertension; congenital heart defect; aortic stenosis; valve disease;					
XX	KW	atrial septal defect; atrioventricular canal defect; ductus arteriosus;					
XX	KW	pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;					
XX	KW	tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;					
XX	KW	obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;					
XX	KW	Parkinson's disease; immune disorder; haematopoietic disorder;					
XX	KW	haemophilia; hypercoagulation; Crohn's disease; cancer.					
XX	OS	Homo sapiens.					
XX	PN	W0200281498-A2.					
XX	PD	17-OCT-2002.					
XX	PF	03-APR-2002; 2002WO-US010780.					
XX	PR	03-APR-2001; 2001US-0281086P.					
XX	PR	03-APR-2001; 2001US-0281136P.					
XX	PR	05-APR-2001; 2001US-0281863P.					
XX	PR	05-APR-2001; 2001US-0281906P.					
XX	PR	06-APR-2001; 2001US-0282020P.					
XX	PR	10-APR-2001; 2001US-028230P.					
XX	PR	10-APR-2001; 2001US-0282934P.					
XX	PR	12-APR-2001; 2001US-0283512P.					
XX	PR	13-APR-2001; 2001US-0283710P.					
XX	PR	17-APR-2001; 2001US-0284234P.					
XX	PR	19-APR-2001; 2001US-0285325P.					
XX	PR	20-APR-2001; 2001US-0285381P.					
XX	PR	20-APR-2001; 2001US-0285609P.					
XX	PR	23-APR-2001; 2001US-0285748P.					

DT	11-AUG-2003 (first entry)	
XX	Human ERK1 protein.	
XX	DARPP-32; psychotic disorder; intracellular signaling protein; dopamine- and cAMP-regulated phosphoprotein; ERK1; ERK2; CREB; extracellular signal-regulated protein kinase; phosphorylation; cAMP-response element binding protein; schizophrenia; human.	
XX	Homo sapiens.	
XX	WO2003021225-A2.	
PN	13-MAR-2003.	
XX	03-SEP-2002; 2002WO-US027802.	
XX	31-AUG-2001; 2001US-0316338P.	
PR	(UYRQ) UNIV ROCKEFELLER.	
PA	(KARO-) KAROLINSKA INST.	
XX	Greengard P, Pisone G;	
PI	WPI; 2003-300912/29.	
XX	Identifying agent to be tested for ability to treat psychotic disorder, by contacting cells/tissues with candidate drug, determining phosphorylation levels of intracellular signaling proteins DARPP-32, ERK1, ERK2, CREB.	
XX	Disclosure; Page 77-78; 79pp; English.	
XX	The invention relates to identifying an agent to be tested for ability to treat psychotic disorder in patient. The method involves contacting cells or tissues with a candidate drug, and determining levels of phosphorylation of intracellular signaling proteins DARPP-32 (dopamine- and cAMP-regulated phosphoprotein), ERK1 and ERK2 (extracellular signal-regulated protein kinases 1 and 2), and CREB (cAMP-response element binding protein). The method is useful for identifying an agent to be tested for an ability to treat a psychotic disorder such as schizophrenia in a patient in need of such treatment. The present sequence represents a human ERK1 protein	
XX	Sequence 379 AA;	
SQ		
Query Match	98.9%; Score 1859; DB 6; Length 379;	
Best Local Similarity	94.7%; Pred. No. 1.1e-184;	
Matches 359; Conservative	0; Mismatches 0; Indels 20; Gaps 1;	
QY	1 MAAAAAGGGGCGGPRTEGVPGEVEMVKQPPFDVGPRTQLOVIGGAYGMVSSAY 60	
DB	1 MAAAAAGGGGCGGPRTEGVPGEVEMVKQPPFDVGPRTQLOVIGGAYGMVSSAY 60	
QY	61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120	
DB	61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120	
QY	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	
DB	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	
QY	181 KICDFGLARTADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEMLS 240	
DB	181 KICDFGLARTADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEMLS 240	
QY	241 NRPIFGKHVLDQLNHLIGLSPSQEDLNCIINNKARNYLQSLPSKTKVAKLPPKSD 300	
DB	241 NRPIFGKHVLDQLNHLIGLSPSQEDLNCIINNKARNYLQSLPSKTKVAKLPPKSD 300	
QY	301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLKPERL 340	
DB	301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLKPERL 340	
QY	341 KELIFQETARFQGVLEAP 359	
DB	361 KELIFQETARFQGVLEAP 379	
RESULT 7		
ADP45050		
ID	ADP45050 standard; protein; 379 AA.	
XX	ADP45050;	
XX	12-FEB-2004 (first entry)	
DE	Human kinase ERK1.	
XX	Human; protein kinase; enzyme; inhibitor; ERK1.	
OS	Homo sapiens.	
XX	WO2003081210-A2.	
PN	02-OCT-2003.	
XX	20-MAR-2003; 2003WO-US008725.	
XX	21-MAR-2002; 2002US-0366892P.	
PA	(SUNE-) SUNESIS PHARM INC.	
XX	Prescott JC, Braisted A;	
PI	WPI; 2003-865136/80.	
XX	Identifying ligand binding to inactive conformation of target protein kinase (T) comprises contacting the conformation modified (T) which contains reactive group at binding site, with ligands and detecting kinase-ligand conjugate formation.	
XX	Disclosure; SEQ ID NO 19; 260pp; English.	
XX	The present invention relates to a method for identifying a ligand (L), which binds to an inactive conformation of target protein kinase (T). The method involves contacting inactive conformation of (T), which contains or is modified to contain a reactive group at or near a binding site of interest, with one or more ligand candidates capable of covalently bonding to the reactive group thus forming a kinase-(L) conjugate (C). The method is useful for identifying protein kinase inhibitors that preferentially bind to inactive conformation of a target protein kinase. The present sequence is a protein kinase which may be modified via an amino acid substitution, for use in the method of the invention.	
XX	Sequence 379 AA;	
SQ		
Query Match	98.9%; Score 1859; DB 7; Length 379;	
Best Local Similarity	94.7%; Pred. No. 1.1e-184;	
Matches 359; Conservative	0; Mismatches 0; Indels 20; Gaps 1;	
QY	1 MAAAAAGGGGCGGPRTEGVPGEVEMVKQPPFDVGPRTQLOVIGGAYGMVSSAY 60	
DB	1 MAAAAAGGGGCGGPRTEGVPGEVEMVKQPPFDVGPRTQLOVIGGAYGMVSSAY 60	
QY	61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120	
DB	61 DHVKTRVAIKKISPFHEQTYCQRTLEIQILFRHENVIGIRDILRASTLEAMRDVYI 120	
QY	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	
DB	121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDL 180	
QY	181 KICDFGLARTADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEMLS 240	
DB	181 KICDFGLARTADPEHDHTGFLTEVATRWTRAPIMLNSKGYTKSIDWSVGCILAEMLS 240	
QY	241 NRPIFGKHVLDQLNHLIGLSPSQEDLNCIINNKARNYLQSLPSKTKVAKLPPKSD 300	
DB	241 NRPIFGKHVLDQLNHLIGLSPSQEDLNCIINNKARNYLQSLPSKTKVAKLPPKSD 300	
QY	301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLKPERL 340	
DB	301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLKPERL 340	

QY 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
DB 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDFPVAEPPFTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 8
ADH59632
ID ADH59632 standard; protein; 379 AA.
XX
AC ADH59632;
XX
DE 25-MAR-2004 (first entry)
DT
DE Erk1 protein.
XX
KW cardiac disease; Raf-1; MEK1; Cardiant; telethonin; &agr-actinin;
KW hypertrophy; MEK2; Erk1; Erk2.

XX Homo sapiens.
XX WO2003025205-A2.
XX
PD 27-MAR-2003.
XX
PF 18-SEP-2002; 2002WO-EP010489.
XX
PR 19-SEP-2001; 2001US-0323566P.
PR 24-SEP-2001; 2001US-0324625P.
XX
PA (MEDI-) MEDIGENE AG.

XX Nave B, Roenicke V, Leclair S, Funk M, Reuner B, Brinkmann K;
PI Henkel T;
XX
DR WPI: 2003-371821/35.
DR N-PSDB; ADH59631.

XX Identifying and/or obtaining a compound useful for preventing or treating
PT cardiac diseases, particularly congestive heart failure, comprises
PT quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
PT on the compound.

PS Claim 33; SEQ ID NO 8; 41pp; English.

XX The present invention relates to identifying and obtaining a compound
CC useful in the prevention or treatment of cardiac diseases, comprising
CC quantification of the activity of Raf-1, MEK1/2 and/or Erk1/2 depending
CC on the compound. The method is useful for identifying and/or obtaining
CC compounds that may be used in the prevention or treatment of cardiac
CC diseases, particularly congestive heart failure. The compound or protein
CC is also used for the preparation of a pharmaceutical composition for
CC prevention or treatment of a disease related to hypertrophy or impaired
CC or increased activation of telethonin (T-Cap), &agr;-actinin, MHC, actin,
CC titin, myomesin, nebulin, tropomyosin, troponin, Erk1/2 and/or MLCK.
CC The protein or antibody which specifically recognizes the
CC activated/phosphorylated form of the above polypeptide, is used for the
CC preparation of a composition for diagnosing a disease or a predisposition
CC for a disease related to hypertrophy or related to impaired or increased
CC activation of Raf-1, MEK1/2 and/or Erk1/2. The present sequence
CC represents Erk1.

XX Sequence 379 AA;

XX Query Match 98.9%; Score 1859; DB 7; Length 379;

Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAOQGGGEPRTTEGVGPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
DB 1 MAAAAOQGGGEPRTTEGVGPGVEVMVKGPFDVGPRTYQLOYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTTCORTLREIQILLRPHENVIGIRDIRASTLEAMRDVVI 120
DB 61 DHVRKTRVAIKKISPEHQTTCORTLREIQILLRPHENVIGIRDIRASTLEAMRDVVI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQIILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICVFLYQIILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVCILAEMLS 240
QY 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
DB 241 NRPIPGKHVLDQLNHILGILGSPSOEDLNCIINMKARNYLOSLSKTKVAKLFPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDFPVAEPPFTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379

RESULT 9
ADO22522
ID ADO22522 standard; protein; 379 AA.
XX
AC ADO22522;
DT
DT 12-AUG-2004 (first entry)
XX
DE Biochemical pathway-related human ERK protein SeqID8.

XX biochemical pathway; mammalian glioma tumour;
KW epidermal growth factor receptor; EGFR; mTOR; polypeptide inhibitor;
KW cancer; ERK; human.

OS Homo sapiens.
XX
PN WO2004044218-A2.
XX
PD 27-MAY-2004.

PF 05-NOV-2003; 2003WO-US035115.
PR 05-NOV-2002; 2002US-0423777P.
XX
PA (REGC) UNIV CALIFORNIA.
PA (CELL-) CELL SIGNALING TECHNOLOGY INC.

XX Mischel PS, Sawyers CL, Smith BL, Crosby K;
XX WPI; 2004-411736/38.

XX Use of biochemical pathways associated with glioblastoma for, e.g.
PT identifying a mammalian glioma tumor that is likely to respond to an
PT epidermal growth factor receptor polypeptide inhibitor or an mTOR
PT polypeptide inhibitor.

PS Claim 1; SEQ ID NO 8; 66pp; English.

XX This invention relates to a novel use of biochemical pathways for
CC identifying a mammalian glioma tumor that is likely to respond to an
CC epidermal growth factor receptor (EGFR) polypeptide inhibitor or an mTOR

CC polypeptide inhibitor, or identifying a mammalian glioma tumour that does
CC not express or expresses a PTEN polypeptide and which is likely to
CC respond or not likely to respond to an inhibitor of mTOR polypeptide
CC activity or inhibitor of EGFR polypeptide activity, respectively. The
CC biochemical pathways are, in particular, dysregulated in pathologies such
CC as cancer. The present sequence is that of the human ERK protein which
CC may be used in the method of the invention.

XX SQ Sequence 379 AA;
Query Match 98.9%; Score 1859; DB 8; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
Qy 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFDPVGRYQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFDPVGRYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVTRKTRVAIKKISPPFHEQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVTRKTRVAIKKISPPFHEQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDINSGVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDINSGVGCILAEMLS 240
Qy 241 NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Db 241 NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAPEPFFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYYDPTDFVABEPTFFAMELDDLPKERL 360
Qy 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 10
ADQ37846
ID ADQ37846 standard; protein; 379 AA.
XX AC ADQ37846;
XX AC
XX DT
XX DE 09-SEP-2004 (first entry)
XX DE Human phosphorylated ERK protein, target for HER-2 therapy Seqid 3.
XX KW human; HER-2 directed therapy; tumour; insulin growth factor receptor;
XX KW IGF1R; epidermal growth factor receptor; EGFR; S6 ribosomal protein; AKT;
XX KW NDF; ERK; cancer therapy; predictive biomarker; HER-2/neu.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Modified-site 202 /note= "Optionally a phosphorylated serine (if threonine
XX FT at 204 is unmodified)"
XX FT Modified-site 204 /note= "Optionally a phosphorylated threonine (if serine
XX FT at 202 is unmodified)"
XX PN WO2004053497-A2.
XX PD 24-JUN-2004.
XX PF 11-DEC-2003; 2003WO-US039770.
XX XX

PR 11-DEC-2002; 2002US-0432942P.
XX (VENT-) VENTANA MEDICAL SYSTEMS INC.
PA (CELL-) CELL SIGNALING TECHNOLOGY INC.
XX
XX Bacus SS, Smith BL;
XX WPI; 2004-507069/48.
XX
XX Identifying a mammalian tumor, useful for predicting a response to HER2-
PT directed therapy, comprises assaying a sample to detect a pattern of
PT expression, phosphorylation or both, of one or more polypeptides.
XX
XX Claim 35; SEQ ID NO 3; 49pp; English.
XX
XX This invention relates to a novel method for determining or predicting
CC the response of a patient to HER-2 directed therapy. Specifically, it
CC refers to analysing a mammalian tumour in order to detect a pattern of
CC expression and/or phosphorylation of a protein taken from the group
CC including insulin growth factor receptor (IGFR) polypeptide, epidermal
CC growth factor receptor (EGFR), phosphorylated NDF or phosphorylated ERK protein. The
CC present invention describes characterising a mammalian tumour's
CC responsiveness to an HER-2 therapy and hence an individual's response to
CC this cancer therapy, using an immunologically specific antibody directed
CC against one of the aforementioned proteins. Furthermore, it provides
CC predictive biomarkers that can be used to assess the efficacy of
CC therapeutic agents targeted to HER-2/neu. This polypeptide sequence is
CC the human phosphorylated ERK protein of the invention.

XX SQ Sequence 379 AA;
Query Match 98.9%; Score 1859; DB 8; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
Qy 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFDPVGRYQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRRTGEGVPGVEVMVKQPFDPVGRYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVTRKTRVAIKKISPPFHEQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVTRKTRVAIKKISPPFHEQTYCQRTLEIQLLRFRHENVIGIRDILRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDINSGVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDINSGVGCILAEMLS 240
Qy 241 NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Db 241 NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAPEPFFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYYDPTDFVABEPTFFAMELDDLPKERL 360
Qy 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 11
ADV42070
ID ADV42070 standard; protein; 379 AA.
XX AC ADV42070;
XX AC
XX DT 24-FEB-2005 (first entry)
XX XX

QY 1 MAAAAAQQGGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOVIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
QY 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180
DB 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLSLPSKTKVAVAKLPKSD 300
DB 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLSLPSKTKVAVAKLPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPEPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEERALAHPYLEQYVDPTEPVAEPEPTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379
RESULT 13
ID ADZ65036 standard; protein; 379 AA.
AC ADZ65036;
XX ADZ65036;
DT 14-JUL-2005 (first entry)
DE Mitogen activated protein kinase, MAPK-3, SEQ ID 2.
KW MAP kinase; Mitogen-activated protein kinase; enzyme; cancer.
OS Homo sapiens.
PN US2005095657-A1.
PD 05-MAY-2005.
XX 21-SEP-2004; 2004US-00945684.
PF 11-OCT-2002; 2002US-0418038P.
PR 14-OCT-2003; 2003WO-US032248.
XX (ARBI/) ARBISER J L.
PA (COHE/) COHEN C.
XX Arbisser JL, Cohen C;
PI WPI; 2005-354474/36.
DR REFSEQ; XP_055766.3.
XX Detecting phosphorylated mitogen activated protein kinase (P-MAPK), by
PT contacting sample with antibody having affinity for P-MAPK, detecting
PT antibody/P-MAPK complex to detect P-MAPK, and expression of P-MAPK
PT indicates cancer.
XX Claim 3; SEQ ID NO 2; 23pp; English.
PS The present invention relates to a method (M1) for detecting
XX Phosphorylated Mitogen Activated Protein Kinase (P-MAPK; ADZ65035-
CC ADZ65042), by providing sample, contacting the sample with at least one
CC antibody having affinity for the phosphorylated portion of P-MAPK, and
CC detecting antibody/P-MAPK complex, in which the presence of antibody/P-
CC MAPK complex indicates that P-MAPK is present in sample, and where

CC constitutive expression of P-MAPK is indicative of cancer.
XX
SQ Sequence 379 AA;
Query Match 98.9%; Score 1859; DB 9; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.1e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAAAQQGGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOVIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHQYTCORTLREIQILLRFRHENVIGIRDLIRASTLEAMRDVYI 120
QY 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180
DB 121 VDQIMETDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLHRLDKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLSLPSKTKVAVAKLPKSD 300
DB 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIIMKARNYLSLPSKTKVAVAKLPKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPEPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNRITVEERALAHPYLEQYVDPTEPVAEPEPTFAMELDDLPKRL 360
QY 341 KELIFQETARFQGVLEAP 359
DB 361 KELIFQETARFQGVLEAP 379
RESULT 14
AAW40540
ID AAW40540 standard; protein; 403 AA.
XX AAW40540;
AC AAW40540;
DT 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 5471.
XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
OS WO200153312-A1.
PN 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
PF 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-0052317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSB-) HYSBQ INC.
PA

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI59696.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
PT
XX
XX Example 2; SEQ ID NO 5471; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAW38442-AAW42213) with nootropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
XX Sequence 403 AA;
XX
Query Match 98.9%; Score 1859; DB 4; Length 403;
Best Local Similarity 94.7%; Pred. No. 1.2e-184;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 1 MAAAQAQGGGGEPRTGEGVGPVGEVEMVKGQPDFVGPRTQLQYIGEGAYGMVSSAY 60
Db 25 MAAAQAQGGGGEPRTGEGVGPVGEVEMVKGQPDFVGPRTQLQYIGEGAYGMVSSAY 84
QY 61 DHVRKTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 85 DHVRKTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDLRASTLEAMRDVYI 144
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Db 145 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 204
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 205 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 264
QY 241 NRPIFGKHVLDQNLHILGILGPSQEDLNCIINMKARNYLOSLPDKTKVAKLPFKSD 300
Db 265 NRPIFGKHVLDQNLHILGILGPSQEDLNCIINMKARNYLOSLPDKTKVAKLPFKSD 324
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPETFAWELDDLPKRL 340
Db 325 SKALDLDRLMTFNPKNRITVEALAHPLYEQYDPTDPSVAEPPETFAWELDDLPKRL 384
QY 341 KELIFOETARFQGVLEAP 359
Db 385 KELIFOETARFQGVLEAP 403
RESULT 15
AAV42413
ID AAV42413 standard; protein; 379 AA.
XX
AC AAV42413;
XX
DT 02-DEC-1999 (first entry)
XX
XX Extracellular signal Regulated Kinase (ERK)1 mutant.

XX
KW mitogen activated protein; MAP kinase; apoptosis; cancer; inflammation;
KW intracellular signal transduction pathway; inhibitor; wildtype; ERK;
KW extracellular signal regulated kinase; pyridinyl-imidazole.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 122
FT /note= "Glu can be substituted by an amino acid with a
FT small side chain, for example alanine or threonine"
XX
XX W09942592-A1.
PN
XX 26-AUG-1999.
PD
XX 16-FEB-1999; 99WO-US003181.
PF
XX 18-FEB-1999; 98US-00025580.
PR
XX (VERT-) VERTEX PHARM INC.
PA
XX Su MS, Fox E, Wilson KP, Germann UA;
PI WPI; 1999-540310/45.
XX
XX Method of designing Ser/Thr or Tyr kinase inhibitor useful for treating,
PT example breast cancer, restenosis, asthma or hypertension.
PT
XX Example 6; Page 50-51; 71pp; English.
PS
XX This is the amino acid sequence of ERK1. Substitution of glutamine with
CC an amino acid with a small side chain, such as alanine and threonine will
CC create a mutant kinase which can theoretically bind to a pyridinyl-
CC imidazole inhibitor of p38. (MAP) 1 kinase. MAP 1 kinases mediate
CC intracellular signal transduction pathways and so have a role in many
CC diverse human diseases. For example, kinases have been implicated in cell
CC entry into apoptosis, cancer, Alzheimer's disease, angiotensin II and
CC hematopoietic cytokine receptor signal transduction, oncoprotein
CC signalling and mitosis, inflammation and infection, etc. Members of the
CC MAP kinase family share sequence similarity and conserved structural
CC domains, and include the extracellular-signal regulated kinases (ERKs),
CC Jun N-terminal kinases (JNKs) and p38 kinases. The invention relates to
CC methods for designing inhibitors of serine/threonine kinases,
CC particularly MAP kinases, and tyrosine kinases through the use of ATP-
CC binding site mutants of these kinases. The methods of this invention take
CC advantage of the fact that the mutant kinases are capable of binding
CC inhibitory compounds of other kinases with greater affinity than the
CC corresponding wild-type kinase
XX
XX Sequence 379 AA;
XX
Query Match 98.6%; Score 1853; DB 2; Length 379;
Best Local Similarity 94.5%; Pred. No. 4.6e-184;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;
QY 1 MAAAQAQGGGGEPRTGEGVGPVGEVEMVKGQPDFVGPRTQLQYIGEGAYGMVSSAY 60
Db 1 MAAAQAQGGGGEPRTGEGVGPVGEVEMVKGQPDFVGPRTQLQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPFHQYTCQRTLEIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILGPSQEDLNCIINMKARNYLOSLPDKTKVAKLPFKSD 300

Db	241	NREIFEGKHYLDQNLHILGILGSPQEDLNCIIINMKARNYLQSLPSKTKVAMAKLFPKSD	300
Qy	301	SKALDLDRLMTENPNKRIT-----VAEPEPTFAMELDDLPKERL	340
Db	301	SKALDLDRLMTENPNKRITVEEALAHPLYEQYYDPTDBEVAEPEPTFAMELDDLPKERL	360
Qy	341	KELIFOETARFPQGVLEAP	359
Db	361	KELIFOETARFPQGVLEAP	379

Search completed: February 6, 2006, 15:29:02
Job time : 287.562 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:29:28 ; Search time 26.8991 Seconds
(without alignments)
1284.124 Million cell updates/sec

Title: US-10-623-108-4

Perfect score: 1879

Sequence: 1 MAAAAAQGGGGPRRTGV.....LKELIFQETARFGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_80.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1853	98.6	379	1 A48082	MAP kinase 3 (EC 2
2	1784	94.9	380	2 JCI451	Ca2+/calmodulin-de
3	1783	94.9	377	2 S28184	Ca2+/calmodulin-de
4	1767	94.0	369	2 A60041	Ca2+/calmodulin-de
5	1550.5	82.5	392	2 JMW052	extracellular sign
6	1538.5	81.9	360	1 JQ1400	MAP kinase 1 (EC 2
7	1538.5	81.9	360	2 S25011	protein kinase ERK
8	1528.5	81.3	358	2 S16444	mitogen-activated
9	1528.5	81.3	358	2 A40033	protein kinase (EC
10	1516	80.7	361	2 A39754	mitogen-activated
11	1508	80.3	369	2 JMW053	extracellular sign
12	1364.5	72.6	376	2 A46036	extracellular sign
13	1317	70.1	376	2 A36978	MAP kinase mpk-1 (
14	1317	70.1	444	2 A36977	MAP kinase sur-1 (
15	932	49.6	356	2 T51944	pathogenicity MAP
16	918.5	48.9	361	2 T51943	mitogen-activated
17	884	47.0	372	2 S15663	protein kinase (EC
18	865	46.0	415	2 A56042	mitogen-activated
19	859.5	45.7	362	2 A47211	protein kinase ERK
20	837.5	44.6	387	2 S48123	mitogen-activated
21	827	44.0	394	2 S33635	mitogen-activated
22	826.5	44.0	371	2 S51320	mitogen-activated
23	822	43.7	376	2 F96619	protein T3016.13
24	822	43.7	395	2 S40472	mitogen-activated
25	817	43.5	393	2 S51321	mitogen-activated
26	814	43.3	372	2 S39559	mitogen-activated
27	814	43.3	384	2 S2989	mitogen-activated
28	813	43.3	376	2 S40470	mitogen-activated
29	807.5	43.0	815	2 B56708	extracellular sign

RESULT 1

A48082

MAP kinase 3 (EC 2.7.1.1) - human

N;Alternate names: extracellular signal-regulated kinase 1 (ERK1); mitogen-activated prot

N;Contains: protein kinase (EC 2.7.1.37)

C;Species: Homo sapiens (man)

C;Date: 04-Sep-1998 #sequence revision 04-Sep-1998 #text_change 09-Jul-2004

C;Accession: A48082; PQ0270; S23428; S15519; S21579

R;Character: D.L.; Mordret, G.; Harder, K.W.; Jirik, F.; Pelech, S.L.

Mol. Cell. Biol. 13, 4679-4690, 1993

A;Title: Molecular cloning, expression, and characterization of the human mitogen-activat

A;Reference number: A48082; MUID:93330262; PMID:7687743

A;Accession: A48082

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-379 <CHA>

A;Cross-references: UNIPROT:P27361; UNIPARC:UPI0000035BE2; EMBL:X60188; NID:g31220; PIDN

A;Experimental source: hepatoma cell line HEP G2

A;Note: authors translated the codon AGC for residue 174 as Ile

R;Owaki, H.; Makar, R.; Boulton, T.G.; Cobb, M.H.; Geppert, T.D.

Biochem. Biophys. Res. Commun. 182, 1416-1422, 1992

A;Title: Extracellular signal-regulated kinases in T cells: characterization of human ERK1

A;Reference number: JQ1400; MUID:92171961; PMID:1540184

A;Accession: PQ0270

A;Molecule type: mRNA

A;Residues: 14-173, 'I', 175-379 <OWA>

A;Cross-references: UNIPARC:UPI000016AB99; GB:M84490; NID:g186695; PIDN:AAA36142.1; PIDN

A;Experimental source: cell line CEM

R;Gonzalez, F.A.; Raden, D.L.; Rigby, M.R.; Davis, R.J.

FEBS Lett. 304, 170-178, 1992

A;Title: Heterogeneous expression of four MAP kinase isoforms in human tissues.

A;Reference number: S23426; MUID:92316223; PMID:1319925

A;Accession: S23428

A;Molecule type: mRNA

A;Residues: 25-173, 'I', 175-379 <GON>

A;Cross-references: UNIPARC:UPI000016A2B4; EMBL:Z11696; NID:g233882; PIDN:CAA77754.1; PID

C;Comment: This enzyme is activated by MAP kinase kinase (see PIR:45100 and PIR:A46723).

C;Genetics:

A;Gene: GDB:PRKM3; ERK1

A;Cross-references: GDB:I35679; OMIM:601795

A;Map position: 16pter-16qter

C;Complex: monomer

C;Function:

A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine

A;Pathway: MAP kinase cascade

C;Superfamily: kinase-related transforming protein; protein kinase homology

C;Keywords: ATP; monomer; phosphoprotein; phosphotransferase; serine/threonine-specific

F;40-330/Domain: protein kinase homology <KIN>

F;48-56/Region: protein kinase ATP-binding motif

F;202/Binding site: phosphate (Thr) (covalent) (by MAP kinase kinase) #status predicted

F;204/Binding site: phosphate (Tyr) (covalent) (by MAP kinase kinase) #status predicted

extracellular sign

mitogen-activated
mitogen-activated
mitogen-activated
probable serine/th
hypothetical prote
protein kinase (EC
mitogen-activated
protein Fl4N23.9 [
mitogen-activated
mitogen-activated
mitogen-activated
mitogen-activated
protein kinase MMK
MAP kinase (ATMPK7
mitogen-activated
hypothetical prote

ALIGNMENTS

Query Match		98.6%;	Score 1853;	DB 1;	Length 379;
Best Local Similarity		94.5%;	Pred. No. 1.1e-77;		
Matches 359;		Conservative	0;	Mismatches 1;	Indels 20; Gaps 1;
QY	1	MAAAAGGGGGPRRTGEGVGPVGEVEMVKQPPFDVGPRTYQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAGGGGGPRRTGEGVGPVGEVEMVKQPPFDVGPRTYQLOYIGEGAYGMVSSAY	60		
QY	61	DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFHENVIGIRDILRASTLEAMRDVYI	120		
DB	61	DHVRKTRVAIKKISPEHQTYCQRTIREIQILLRFHENVIGIRDILRASTLEAMRDVYI	120		
QY	121	VDLMTDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSA NVLHRDLKPSNLLINTTCDL	180		
DB	121	VDLMTDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSA NVLHRDLKPSNLLINTTCDL	180		
QY	181	KICDFGLARIADPEHDHTGFLT EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS	240		
DB	181	KICDFGLARIADPEHDHTGFLT EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLS	240		
QY	241	NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSD	300		
DB	241	NRPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSD	300		
QY	301	SKALDLLDRMLTFNPKNKRIT-----VAEPTTFAMELDDLPERLK	340		
DB	301	SKALDLLDRMLTFNPKNKRITVEALAHPLYEQYVDTDPVABEPTTFAMELDDLPERLK	360		
QY	341	KELIFQETARFQGVLEAP	359		
DB	361	KELIFQETARFQGVLEAP	379		
RESULT 2					
JC1451					
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) - rat					
N/Alternate names: ERK1-MAP kinase; extracellular signal-regulated kinase					
C/Species: Rattus norvegicus (Norway rat)					
C/Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004					
C/Accession: JC1451; A35061; A37140; A40466; S24947					
R/Marquardt, B.; Stabel, S.					
Gene 120, 297-299, 1992					
A/Title: Sequence of a rat cDNA encoding the ERK1-MAP kinase.					
A/Reference number: JC1451; MUID:93013050; PMID:1327976					
A/Accession: JC1451					
A/Molecule type: mRNA					
A/Residues: 1-380 <WAR>					
A/Cross-references: UNIPROT:P21708; UNIPARC:UPI000012F174; EMBL:X65198; NID:g56626; PIDN:AA1123.1					
A/Experimental source: brain					
R/Boulton, T.G.; Yancopoulos, G.D.; Gregory, J.S.; Slaughter, C.; Moomaw, C.; Hsu, J.; C					
Science 249, 64-67, 1990					
A/Title: An insulin-stimulated protein kinase similar to yeast kinases involved in cell					
A/Reference number: A35061; MUID:90312137; PMID:2164259					
A/Accession: A35061					
A/Molecule type: mRNA					
A/Residues: 14-380 <BOU>					
A/Cross-references: UNIPARC:UPI0000145072; GB:M38194; NID:g204051; PIDN:AAA41123.1; PID:					
R/Boulton, T.G.; Gregory, J.S.; Cobb, M.H.					
Biochemistry 30, 278-286, 1991					
A/Title: Purification and properties of extracellular signal-regulated kinase 1, an insu					
A/Reference number: A37140; MUID:91105092; PMID:1846291					
A/Accession: A37140					
A/Status: preliminary					
A/Molecule type: protein					
A/Residues: 43-64;167-178, 'X',180-183, 'X',185 <BO2>					
A/Cross-references: UNIPARC:UPI0000175587; UNIPARC:UPI0000175590					
R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.					
DNA Cell Biol. 10, 505-514, 1991					
A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse a					
A/Reference number: A40466; MUID:91369479; PMID:1716439					
A/Accession: A40466					
A/Status: preliminary					
A/Molecule type: mRNA					

A/Residues: 14-94, 'R',96-380 <DEM>					
A/Cross-references: UNIPARC:UPI00001707CE; GB:U12008; GB:S59509; NID:g515498; PIDN:AAA201					
C/Superfamily: kinase-related transforming protein; protein kinase homology					
C/Keywords: ATP; calmodulin binding; phosphoprotein; phosphotransferase; serine/threonine					
F:41-331/Domain: protein kinase homology <KIN>					
F:49-57/Region: protein kinase ATP-binding motif					
Query Match					
Best Local Similarity 91.8%; Pred. No. 1.5e-74;					
Matches 347; Conservative 1; Mismatches 10; Indels 20; Gaps 1;					
QY	2	AAAAAGGGGGPRRTGEGVGPVGEVEMVKQPPFDVGPRTYQLOYIGEGAYGMVSSAYD	61		
DB	3	AAAAAGGGGGPRRTGAGVVPVPGVEVEMVKQPPFDVGPRTYQLOYIGEGAYGMVSSAYD	62		
QY	62	HVYKTRVAIKKISPEHQTYCQRTIREIQILLRFHENVIGIRDILRASTLEAMRDVYIV	121		
DB	63	HVYKTRVAIKKISPEHQTYCQRTIREIQILLRFHENVIGIRDILRASTLEAMRDVYIV	122		
QY	122	QDLMTDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSA NVLHRDLKPSNLLINTTCDLK	181		
DB	123	QDLMTDLYKLLKSQQLSNDHICYFLYQILRLGKVIHSA NVLHRDLKPSNLLINTTCDLK	182		
QY	182	ICDFGLARIADPEHDHTGFLT EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN	241		
DB	183	ICDFGLARIADPEHDHTGFLT EYVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN	242		
QY	242	RPPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSDS	301		
DB	243	RPPIFGKHYLDQNLHILGILSPSQEDLNCIINMKARNYLOSLPSKTKVAAKLFPKSDS	302		
QY	302	KALDLLDRMLTFNPKNKRIT-----VAEPTTFAMELDDLPERLK	341		
DB	303	KALDLLDRMLTFNPKNKRITVEALAHPLYEQYVDTDPVABEPTTFAMELDDLPERLK	362		
QY	342	ELIFQETARFQGVLEAP	359		
DB	363	ELIFQETARFQGAPEAP	380		
RESULT 3					
S28184					
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) erk-1 - mouse (fragment)					
N/Alternate names: extracellular signal-regulated kinase 1					
C/Species: Mus musculus (house mouse)					
C/Date: 22-Nov-1993 #sequence revision 30-Jan-1998 #text_change 09-Jul-2004					
C/Accession: S28184; B40466; A41371					
R/Tanner, B.; Mueckler, M.					
Biochim. Biophys. Acta 1171, 319-320, 1993					
A/Title: Molecular cloning of a mouse extracellular signal regulated kinase (erk-1).					
A/Reference number: S28184; MUID:93144347; PMID:8424957					
A/Accession: S28184					
A/Molecule type: mRNA					
A/Residues: 1-377 <TAN>					
A/Cross-references: UNIPROT:Q63844; UNIPARC:UPI000017558C; EMBL:Z14249					
R/De Miguel, C.; Kligman, D.; Patel, J.; Detera-Wadleigh, S.D.					
DNA Cell Biol. 10, 505-514, 1991					
A/Title: Molecular analysis of microtubule-associated protein-2 kinase cDNA from mouse a					
A/Reference number: A40466; MUID:91369479; PMID:1716439					
A/Accession: B40466					
A/Molecule type: mRNA					
A/Residues: 74-377 <DEM>					
A/Cross-references: UNIPARC:UPI0000023599; GB:S59517					
R/Crews, C.W.; Alessandrini, A.A.; Erikson, R.L.					
Proc. Natl. Acad. Sci. U.S.A. 88, 8845-8849, 1991					
A/Title: Mouse Erk-1 gene product is a serine/threonine protein kinase that has the poter					
A/Reference number: A41371; MUID:92020947; PMID:1717989					
A/Accession: A41371					
A/Molecule type: mRNA					
A/Residues: 7-16 <CRE>					
A/Cross-references: UNIPARC:UPI0000170C57; GB:S58470; NID:g236372; PIDN:AAB19973.1; PID:g					
C/Superfamily: kinase-related transforming protein; protein kinase homology					
C/Keywords: ATP; calmodulin binding; phosphotransferase; protein kinase					

R;Payne, D.M.; Rossomando, A.J.; Martino, P.; Erickson, A.K.; Her, J.H.; Shabanowitz, J. EMBO J. 10, 885-892, 1991

A;Title: Identification of the regulatory phosphorylation sites in pp42/mitogen-activated protein kinase

A;Reference number: S15015; MUID:91184134; PMID:1849075

A;Accession: S15015

A;Status: preliminary

A;Molecule type: protein

A;Residues: 171-181, 'X', 183-189 <P>

A;Cross-references: UNIPARC:UPI0000175588

R;Ershler, M.A.; Nagorsky, T.V.; Vasser, J.W.M.; Belyavsky, A.V. Gene 124, 305-306, 1993

A;Title: Novel CDC2-related protein kinases produced in murine hematopoietic stem cells

A;Reference number: PN0479; MUID:93185941; PMID:8444355

A;Accession: PN0480

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 145-195 <ER>

A;Cross-references: UNIPARC:UPI0000175589

C;Superfamily: kinase-related transforming protein; protein kinase homology

C;Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine kinase

F;21-311/Domain: protein kinase homology <KIN>

F;29-37/Region: protein kinase ATP-binding motif

F;183/Binding site: phosphate (Thr) (covalent) #status experimental

F;185/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 81.3%; Score 1528.5; DB 2; Length 358;
Best Local Similarity 78.6%; Pred. No. 5.4e-63;
Matches 294; Conservative 19; Mismatches 22; Indels 39; Gaps 2;

Qy	1	MAAAAAAGGGGEPRTTEGVGPGVGEVMVGQPDVGPRTYQLQYIGGAYGMVSSAY	60
Db	1	MAAAAAAG-----PEMVRGQVFDVGPRTYNLSYIGGAYGMVCSAY	41
Qy	61	DHVKTTRVAIKISPPFHQYTCQRTREIOILLRFHENVIGIRDILRASTLEAMRDVYI	120
Db	42	DNLNKRVAIKISPPFHQYTCQRTREIKILLRFHENVIGINDIRATFIQMDVYI	101
Qy	121	VQDLMETDLYKLLKSQSLSDNHCHYFLYQILRLGLKYTHSANVLHRLDKPSNLLINTTCDL	180
Db	102	VQDLMETDLYKLLKTLQHLSDNHCHYFLYQILRLGLKYTHSANVLHRLDKPSNLLINTTCDL	161
Qy	181	KICDFGLARIADPDHDTGFLTEVATWTRAPIMLNSKGYTKSIDINSVGCILAEMLS	240
Db	162	KICDFGLARVADPDHDTGFLTEVATWTRAPIMLNSKGYTKSIDINSVGCILAEMLS	221
Qy	241	NRPFPGKHVLDQNLHILGILGSPQSDLNLCINMKARNYLQSLPSKTKYAAKLPPKSD	300
Db	222	NRPFPGKHVLDQNLHILGILGSPQSDLNLCINMKARNYLLSLPHKKNYPMNRLFPNAD	281
Qy	301	SKALDLDRLMTFNPKNRITV-----ABEPTTFAMELDDLPKERL	340
Db	282	SKALDLDKMLTFNPHKRIVEQALAHPLYEQYYDPSDEPIAEAPFKFDMELDLPKEKL	341
Qy	341	KELIFQETARFPQG 354	
Db	342	KELIFEETARFPQG 355	

RESULT 9

A40033

protein kinase (EC 2.7.1.37) ERK2 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004

C;Accession: A40033

R;Boulton, T.G.; Nye, S.H.; Robbins, D.J.; Ip, N.Y.; Radziejewska, E.; Morgenbesser, S.D. Cell 65, 663-675, 1991

A;Title: ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated upon mitogenic stimulation

A;Reference number: A40033; MUID:91233502; PMID:2032290

A;Accession: A40033

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-358 <BOU>

A;Cross-references: UNIPROT:P27703; UNIPARC:UPI000000303FEA; GB:M64300; NID:g204055; PMID:

Db 1 MAAAAASSPFGGP-----EMVRQAFDVGPRYTNLSYIGEGAYGMVCSAH 46
QY 61 DHVKTTRVAIKKISPFHEQTYCORTLREIQILFRHENVIGIRDILRASTLEAMRDVYI 120
Db 47 CNINKVRVAIKKISPFHEQTYCORTLREIKILLRFKHENIIGINDIIRAPTIQMKDVYI 106
QY 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 107 VQDLMETDLYKLLKTOHLNSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 166
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGTYKSIDIWSVGCILAEMLS 240
Db 167 KICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLNSKGTYKSIDIWSVGCILAEMLS 226
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVNAKLPFKSD 300
Db 227 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLLSLPHKNKVPWNLFFPNAD 286
QY 301 SKALDILDRMLTFNPKRI-----TVAEPPFTFAMELDDLPKERL 340
Db 287 PKALDILDRMLTFNPKRIEVEAALAHPLYEQYDPSDFVABAPLKFEMLDDLPKETL 346
QY 341 KELIFOETARFQPG 354
Db 347 KELIFETARFQPG 360

RESULT 11
JW0053
extracellular signal-regulated kinase (EC 2.7.-.-) 2 - common carp
N;Alternate names: CERK2
C;Species: Cyprinus carpio (common carp)
C;Date: 13-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
C;Accession: JW0053
R;Biochem. 123, 1031-1035, 1998
J. Biochem. 123, 1031-1035, 1998
A;Title: Existence of two isoforms of extracellular signal-regulated kinase in fish.
A;Reference number: JW0052; MUID:98269030; PMID:96603989
A;Molecule type: mRNA
A;Residues: 1-369 <HAS>
A;Cross-references: UNIPROT:O7LZH5; UNIPARC:UPI0000175597
C;Comment: This enzyme is the key enzyme in the intracellular signal transduction pathway
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase
F;32-322/Domain: protein kinase homology <KIN>

Query Match 80.3%; Score 1508; DB 2; Length 369;
Best Local Similarity 77.7%; Pred. No. 4.7e-62;
Matches 290; Conservative 21; Mismatches 32; Indels 30; Gaps 3;

QY 2 AAAAAQGGGGEPRTTEGVGPGVGEVMEVKGQFDVGPRYTOLOVIGEGAYGMVSSAYD 61
Db 4 AAVSAPAGGGPNP-----GSG-----AEMVRGQAFDVGPRYTNLSYIGEGAYGMVCSAYK 53
QY 62 HVKTRVAIKKISPFHEQTYCORTLREIQILFRHENVIGIRDILRASTLEAMRDVYI 121
Db 54 RDNKVRVAIKKISPFHEQTYCORTLREIKILLRFKHENIIGINDIIRTTIDQMDVYI 113
QY 122 QDLMETDLYKLLKSQQLSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLK 181
Db 114 QDLMETDLYKLLKTOHLNSNDHICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLK 173
QY 182 ICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGTYKSIDIWSVGCILAEMLSN 241
Db 174 ICDGLARVADPDHDTGFLTEYVATRWYRAPEIMLNSKGTYKSIDIWSVGCILAEMLSN 233
QY 242 RPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVNAKLPFKSDS 301
Db 234 RPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLLSLPHKNKVPWNLFFPNADP 293
QY 302 KALDILDRMLTFNPKRI-----TVAEPPFTFAMELDDLPKERLK 341

Db 294 KALDILDRMLTFNPKRIEVEAALAHPLYEQYDPTDEPVAEAPFKFDMELDDLPKETLK 353
QY 342 ELIFOETARFQPG 354
Db 354 ELIFETARFQPG 366

RESULT 12
A46036
extracellular signal-regulated kinase (EC 2.7.1.1-) DmERK-A - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46036
R;Biggs III, W.H.; Zipursky, S.L.
Proc. Natl. Acad. Sci. U.S.A. 89, 6295-6299, 1992
A;Title: Primary structure, expression, and signal-dependent tyrosine phosphorylation of
A;Reference number: A46036; MUID:92335284; PMID:1378625
A;Accession: A46036
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-376 <BIG>
A;Cross-references: UNIPROT:P40417; UNIPARC:UPI0000000FB3; GB:M95124; NID:G485754; PIDN:I
C;Note: sequence extracted from NCBI backbone (NCBIN:108684, NCBI:P:108685)
C;Genetics:
A;Gene: FlyBase:rl
A;Cross-references: FlyBase:FBgn0003256
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k
F;36-326/Domain: protein kinase homology <KIN>
F;44-52/Region: protein kinase ATP-binding motif

Query Match 72.6%; Score 1364.5; DB 2; Length 376;
Best Local Similarity 74.9%; Pred. No. 1.5e-55;
Matches 262; Conservative 28; Mismatches 39; Indels 21; Gaps 2;

QY 24 VP-GEVMEVKGQFDVGPRYTOLOVIGEGAYGMVSSAYDHVKTRVAIKKISPFHEQTYC 82
Db 19 VPQSNAEVIRGQIFEVGPRYIKLAVIGEGAYGMVVSADDTLTNRQVVAIKKISPFHEQTYC 78
QY 83 QRTLREIQILFRHENVIGIRDILRASTLEAMRDVYIQQDLMETDLYKLLKSQQLSNDH 142
Db 79 QRTLREITLTFRKHENIIDIRLDVDSIDQMDRVYIVQCLMETDLYKLLKTQRLSNDH 138
QY 143 ICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLT 202
Db 139 ICYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLT 198
QY 203 EYVATRWYRAPEIMLNSKGTYKSIDIWSVGCILAEMLSNRPIFGKHYLDQNLHILGILG 262
Db 199 EYVATRWYRAPEIMLNSKGTYKSIDIWSVGCILAEMLSNRPIFGKHYLDQNLHILGVLG 258
QY 263 SPSQEDLNCIINMKARNYLOSLPSKTKVNAKLPFKSDSKALDILDRMLTFNPKRI---- 319
Db 259 SPSRDLCEIINEKARNYLESIPFPNVPWAKLFFPNADALADLLGKMLTFNPKRIPVE 318
QY 320 -----TVAEPPFTFAMELDDLPKERLKELIFOETARFQ 352
Db 319 EALAHPLYEQYDPSDFVABEVPFRINMENDDISRDALKSLIFEETLKF 368

RESULT 13
A36978
MAP kinase mpk-1 (EC 2.7.1.1-) - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: A36978; T22118
R;Lackner, M.R.; Kornfeld, K.; Miller, L.M.; Horvitz, H.R.; Kim, S.K.
Genes Dev. 8, 160-173, 1994
A;Title: A MAP kinase homolog, mpk-1, is involved in ras-mediated induction of vulval cel
A;Reference number: A36978; MUID:94131270; PMID:8299936
A;Accession: A36978
A;Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-376 <LAC>
A:Cross-references: UNIPROT:P39745; UNIPARC:UPI000002B465; GB:U27124; GB:S68854; NID:986
R:Jassal, B.; Smith, A
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19519
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-376 <WIL>
A:Cross-references: UNIPARC:UPI000002B465; EMBL:Z46937; PIDN:CAA87057.1; GSPDB:GN00021;
A:Experimental source: clone F43C1
C:Genetics:
A:Gene: mpk-1
A:Map position: 3
A:Introns: 23/3; 43/2; 85/3; 104/2; 304/2
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:26-285/Domain: protein kinase homology <KIN>
F:34-42/Region: protein kinase ATP-binding motif

Query Match 70.1%; Score 1317; DB 2; Length 376;
Best Local Similarity 72.4%; Pred. No. 2.2e-53;
Matches 246; Conservative 36; Mismatches 36; Indels 20; Gaps 1;

QY 28 VEMVKGQFPDVGPRYTQLOIYIGEGAYGMVSSAYDHRVTRVAKKISPFPHQTYCQRTL 87
DB 14 VEEVHGOLFVAPRYVNLISYIGEGAYGMVASALDTITRDRVAIKKISPFPHQTYCQRTL 73
QY 88 EIQILLRFHENVIGIRDIRASTLEARDVYIVQDLMETDLYKLLKSQQLSNDHICYFL 147
DB 74 EIKILNRFKHENIINQIIRSETVDSLKDIYVQCLMETDLYKLLKTOKLSNDHVCYFL 133
QY 148 YQILRGKLYIHSANVLRDLKPSNLLINTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 207
DB 134 YQILRGKLYIHSANVLRDLKPSNLLINTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 193
QY 208 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDLQNLHILGILGSPSQE 267
DB 194 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDLQNLHILAVVGSFNSA 253
QY 268 DLNCIINMKARNYLOSLPSKTKVAVAKLPKSDSKALDLDRLMTFNPKNRI----- 319
DB 254 DLQCIINDKARSYLISLPHKPKQWARLYPGADPRALDLDKMLTFNPHNRIDIEQALAH 313
QY 320 -----TVAEPEPTFAMELDDLPKELKELIFOE 347
DB 314 PYLEQYVDPGDEPVCBEPFTEMEFDDLPKELKELIWE 353

RESULT 14
A36977
MAP kinase sur-1 (EC 2.7.1.1) - Caenorhabditis elegans
A:Title: Suppression of activated Let-60 ras protein defines a role of Caenorhabditis el
A:Reference number: A36977; MUID:94131269; PMID:8299935
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: A36977
R:Wu, Y.; Han, M.
Genes Dev. 8, 147-159, 1994
A:Title: Suppression of activated Let-60 ras protein defines a role of Caenorhabditis el
A:Reference number: A36977; MUID:94131269; PMID:8299935
A:Accession: A36977
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-444 <WUA>
A:Cross-references: UNIPROT:P39745; UNIPARC:UPI0000164181; GB:U03879; NID:q433356; PIDN:
C:Genetics:
A:Gene: sur-1
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:94-353/Domain: protein kinase homology <KIN>
F:102-110/Region: protein kinase ATP-binding motif

Query Match 70.1%; Score 1317; DB 2; Length 444;

Best Local Similarity 72.4%; Pred. No. 2.5e-53;
Matches 246; Conservative 38; Mismatches 36; Indels 20; Gaps 1;

QY 28 VEMVKGQFPDVGPRYTQLOIYIGEGAYGMVSSAYDHRVTRVAKKISPFPHQTYCQRTL 87
DB 82 VEEVHGOLFVAPRYVNLISYIGEGAYGMVASALDTITRDRVAIKKISPFPHQTYCQRTL 141
QY 88 EIQILLRFHENVIGIRDIRASTLEARDVYIVQDLMETDLYKLLKSQQLSNDHICYFL 147
DB 142 EIKILNRFKHENIINQIIRSETVDSLKDIYVQCLMETDLYKLLKTOKLSNDHVCYFL 201
QY 148 YQILRGKLYIHSANVLRDLKPSNLLINTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 207
DB 202 YQILRGKLYIHSANVLRDLKPSNLLINTTCDLKICDFGLARVTDPTDHTGFLTEYVAT 261
QY 208 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDLQNLHILGILGSPSQE 267
DB 262 RWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDLQNLHILAVVGSFNSA 321
QY 268 DLNCIINMKARNYLOSLPSKTKVAVAKLPKSDSKALDLDRLMTFNPKNRI----- 319
DB 322 DLQCIINDKARSYLISLPHKPKQWARLYPGADPRALDLDKMLTFNPHNRIDIEQALAH 381
QY 320 -----TVAEPEPTFAMELDDLPKELKELIFOE 347
DB 382 PYLEQYVDPGDEPVCBEPFTEMEFDDLPKELKELIWE 421

RESULT 15
T51944
pathogenicity MAP kinase 1 [imported] - Pyricularia grisea
C:Species: Pyricularia grisea
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T51944
R:Xu, J.R.; Hamer, J.E.
Genes Dev. 10, 2696-2706, 1996
A:Title: MAP kinase and CAMP signaling regulate infection structure formation and pathoge
A:Reference number: Z25880; MUID:97102671; PMID:8946911
A:Accession: T51944
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-356 <XUJ>
A:Cross-references: UNIPROT:Q92246; UNIPARC:UPI000006C824; EMBL:U70134; PIDN:AAC49521.2
A:Experimental source: strain Guy11
C:Genetics:
A:Gene: PMK1
A:Introns: 39/2; 232/1
C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 49.6%; Score 932; DB 2; Length 356;
Best Local Similarity 53.4%; Pred. No. 5.8e-36;
Matches 181; Conservative 55; Mismatches 79; Indels 24; Gaps 2;

QY 36 FDVGPRTQLOIYIGEGAYGMVSSAYDHRVTRVAKKISPFPHQTYCQRTLREIQILLRF 95
DB 18 FNVSEQYDIQDVVGEAYGVVCSAIHKPSQKQVAKKITEFDHSMFCLETLREMKLLRYF 77
QY 96 RHENVIGIRDIRASTLEARDVYIVQDLMETDLYKLLKSQQLSNDHICYFLYQILGLK 155
DB 78 NHENISILDIQKPRSVETNEVYLIQELMETDHRVIRVTDLSDDHCQFYIQTALRK 137
QY 156 YIHSANVLRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAPEI 215
DB 138 AMHSANVLRDLKPSNLLNANCDLVKDFGLARSAAQENNSGFMTYVATRWYRAPEI 197
QY 216 MLNSGYTKSIDIWSVGCILAEMLSNRPIFFPGKHLYDLQNLHILGILGSPSQEDLNCIINM 275
DB 198 MLTFKTYKTRADVWSVGCILAEMLSNGKPLFGKDYHHQLTLILDVLGTPTWEDYIGKSR 257
QY 276 KARNYLOSLPSKTKVAVAKLPKSDSKALDLDRLMTFNPKNRI----- 323
DB 258 RAREYINSLFPKKVPPRTLPFKTSDLAADLLEKLLAFNPKRITVEALKHPYLEPYHD 317

QY 324 ---EP-----FTFAMELDDLPKERLKEIFQETAR 350
Db 318 PDDEPTAPIPEEFFDDXHKONLSKEQLXQFIYQEIIR 356

Search completed: February 6, 2006, 15:35:34
Job time : 27.8991 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:20:01 ; Search time 171.223 Seconds
(without alignments)
1479.265 Million cell updates/sec

Title: US-10-623-108-4
Perfect score: 1879
Sequence: 1 MAAAAAQGGGGGPRRTGV.....LKELIFQTARFGVLEAP 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1859	98.9	379	1 MK03 HUMAN	P27361 h mitogen-a
2	1787	95.1	380	1 MK03 MOUSE	Q63844 h mitogen-a
3	1784	94.9	380	1 MK03 RAT	P21708 r mitogen-a
4	1784	94.9	380	2 QAPI78 RAT	Q4piV8 rattus norv
5	1767	94.0	369	2 Q7M0H9 CRIGR	Q7m0h9 cricetus
6	1683	89.6	357	2 Q8NHX1 HUMAN	Q8nhx1 homo sapien
7	1593	84.8	335	2 Q8NHX0 HUMAN	Q8nhx0 homo sapien
8	1561	83.1	323	2 Q7Z3H5 HUMAN	Q7z3h5 homo sapien
9	1551.5	82.6	392	2 Q7ZVK8 BRARE	Q7zvK8 brachydanio
10	1550.5	82.5	391	2 Q9DGR6 BRARE	Q9dgr6 brachydanio
11	1550.5	82.5	392	2 Q7LZH4 CYPCA	Q7lzh4 cyprinus ca
12	1538.5	81.9	360	1 MK01 BOVIN	P46196 bos taurus
13	1538.5	81.9	360	1 MK01 HUMAN	P28482 gallus gall
14	1537	81.8	368	2 Q8UW66 CHICK	Q8uw66 gallus gall
15	1528.5	81.3	358	1 MK01 MOUSE	P63085 mus musculus
16	1528.5	81.3	358	1 MK01 RAT	P63086 rattus norv
17	1527	81.3	358	2 Q8NVR1 PONPY	Q8nvr1 pongo pygma
18	1522	81.0	361	1 MK01 XENLA	P26696 xenopus lae
19	1520	80.9	361	2 Q5D0K1 XENLA	Q5d0k1 xenopus lae
20	1511	80.4	369	2 Q6P023 BRARE	Q6p023 brachydanio
21	1509	80.3	361	2 Q6DFK6 XENLA	Q6dfk6 xenopus lae
22	1508	80.3	369	2 Q7LZH5 CYPCA	Q7lzh5 cyprinus ca
23	1506	80.1	369	2 Q5I811 CYPCA	Q5i811 cyprinus ca
24	1504	80.0	369	2 Q7ZW72 BRARE	Q7zw72 brachydanio
25	1501	79.9	369	2 Q9DGR5 BRARE	Q9dgr5 brachydanio
26	1495	79.6	369	2 Q4VH14 BRARE	Q4vh14 brachydanio
27	1488	79.2	358	2 Q4S9M7 TETNG	Q4s9m7 tetraodon n
28	1488	79.2	366	2 Q4RV04 TETNG	Q4rv04 tetraodon n
29	1445	76.9	351	2 Q16945 AFLCA	Q16945 alysia cal
30	1418	75.5	369	2 Q8IS92 STRPU	Q8is92 strongyloce
31	1399	74.5	350	2 Q4H3L5 CIOIN	Q4h3l5 ciona intes

32	1384.5	73.7	395	2 Q8AV10 PAROL	Q8avi0 parolichthy
33	1372.5	73.0	372	2 Q86N95 MARGL	Q86n95 marthasteri
34	1371.5	73.0	362	2 Q8MXZ5 HALRO	Q8mxz5 halocynthia
35	1364.5	72.6	376	1 ERKA DROME	P40417 drosophila
36	1364.5	72.6	376	2 Q7PL59 DROME	Q7pl59 drosophila
37	1355	72.1	394	2 Q5K4T9 9BILA	Q5k4t9 meloidogyne
38	1323	70.4	374	2 Q6ORP3 CABER	Q6orP3 caenorhabdi
39	1317	70.1	444	1 SURI CAEEL	P39745 caenorhabdi
40	1242.5	66.1	291	2 Q35558 MOUSE	Q35558 mus musculus
41	1163.5	61.9	351	2 Q6PNM5 SCHJA	Q6pnm5 schistosoma
42	1095	58.3	230	2 Q7PVR1 ANOGA	Q7pvr1 anopheles g
43	944	50.2	352	2 Q876L7 LEPMC	Q876l7 leptosphaer
44	944	50.2	352	2 Q96X31 9PLEO	Q96x31 pyrenophora
45	944	50.2	352	2 Q6R3Q3 ALTER	Q6r3q3 alternaria

ALIGNMENTS

RESULT 1	379 AA.
MK03_HUMAN	STANDARD;
ID	379 AA.
AC	P27361;
DT	01-AUG-1992 (Rel. 23, Created)
DT	01-APR-1993 (Rel. 25, Last sequence update)
DT	13-SEP-2005 (Rel. 48, Last annotation update)
DE	Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
DE	signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
DE	(MAP kinase 1) (MAPK 1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
DE	associated protein 2 kinase).
GN	Names=MAPK3; Synonyms=ERK1, PRK3;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC	Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Hepatoma;
RX	MEDLINE=93330262; PubMed=7687743;
RA	Charest D.L., Jirik F., Harder K., Pelech S.L., Mordret G.;
RT	"Molecular cloning, expression, and characterization of the human
RL	mitogen-activated protein kinase p44erk1.";
RN	[2]
RC	TISSUE=Lymph;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U.,
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[3]
RC	NUCLEOTIDE SEQUENCE OF 14-379.
RX	MEDLINE=92171961; PubMed=1540184;
RA	Okaki H., Makar R., Boulton T.G., Cobb M.H., Geppert T.D.;
RT	"Extracellular signal-regulated kinases in T cells: characterization


```

OY 342 ELIFQETARFQGVLEAP 359
Db 363 ELIFQETARFQGVLEAP 380

RESULT 3
MK03 RAT STANDARD; PRT; 380 AA.
AC P21708; Q62686; Q9J1J3;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 3 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase)
DE (MAP kinase 1) (MAPK 1) (p44-ERK1) (ERT2) (p44-MAPK) (Microtubule-
DE associated protein 2 kinase) (MNK1).
DE Names=Mapk3; Synonyms=Erk1, Prkm3;
GN Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=93013050; PubMed=1327976; DOI=10.1016/0378-1119(92)90109-3;
RA Marquardt B., Stabel S.;
RT "Sequence of a rat cDNA encoding the ERK1-MAP kinase.";
RL Gene 120:297-299(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 7-380.
RA Maisonnier P.C., le Beau M.M., Espinosa R. III, Ip N.Y.,
RA Belluscio L., la Monte S.M., Squinco S., Furth M.E., Yancopoulos G.D.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE OF 14-380, AND PARTIAL PROTEIN SEQUENCE.
RX MEDLINE=90312137; PubMed=2164259;
RA Boulton T.G., Yancopoulos G.D., Gregory J.S., Slaughter C., Moomaw C.,
RA Hsu J., Cobb M.H.;
RT "An insulin-stimulated protein kinase similar to yeast kinases
RT involved in cell cycle control.";
RL Science 249:64-67(1990).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 14-380.
RC TISSUE=Brain cortex;
RX MEDLINE=91369479; PubMed=1716439;
RA de Miguel C., Kligman D., Patel J., Detera-Wadleigh S.D.;
RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA
RT from mouse and rat brain.";
RL DNA Cell Biol. 10:505-514(1991).
RN [5]
RP NUCLEOTIDE SEQUENCE (ISOFORM B).
RX MEDLINE=20283605; PubMed=10748187; DOI=10.1074/jbc.M910060199;
RA Yung Y., Yao Z., Hanoch T., Seger R.;
RT "ERK1b, a 46-kDa ERK isoform that is differentially regulated by
RT MEK";
RL J. Biol. Chem. 275:15799-15808(2000).
RN [6]
RP PROTEIN SEQUENCE OF 43-64 AND 167-185, AND CHARACTERIZATION.
RX MEDLINE=91105092; PubMed=1846291;
RA Boulton T.G., Gregory J.S., Cobb M.H.;
RT "Purification and properties of extracellular signal-regulated kinase
RT 1, an insulin-stimulated microtubule-associated protein 2 kinase.";
RL Biochemistry 30:278-286(1991).
RN [7]
RP AUTOPHOSPHORYLATION.
RX MEDLINE=91296777; PubMed=1712480;
RA Seger R., Ahn N.G., Boulton T.G., Yancopoulos G.D., Panayotatos N.,
RA Radziejewska E., Ericsson L., Bratlien R.L., Cobb M.H., Krebs E.G.;
RT "Microtubule-associated protein 2 kinases, ERK1 and ERK2, undergo
RT autophosphorylation on both tyrosine and threonine residues:
RT implications for their mechanism of activation.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:6142-6146(1991).
RN [8]
RP PHOSPHORYLATION OF E1F4BP1.
RX MEDLINE=95025978; PubMed=7939721;
RA Lin T.-A., Kong X., Haystead T.A.J., Pause A., Belsham G.J.,
RA Sonenberg N., Lawrence J.C. Jr.;
RT "PHAS-I as a link between mitogen-activated protein kinase and
RT translation initiation.";
RL Science 266:653-656(1994).
CC -!- FUNCTION: Involved in both the initiation and regulation of
CC meiosis, mitosis, and postmitotic functions in differentiated
CC cells by phosphorylating a number of transcription factors such as
CC ELK-1. Phosphorylates E1F4BP1; required for initiation of
CC translation. Phosphorylates microtubule-associated protein 2
CC (MAP2) (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated and tyrosine phosphorylated in
CC response to insulin and NGF.
CC -!- SUBCELLULAR LOCATION: Nuclear (isoform B).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=P21708-1; Sequence=Displayed;
CC Name=B;
CC IsoId=P21708-2; Sequence=VSP_004830;
CC -!- TISSUE SPECIFICITY: Highest levels within the nervous system,
CC expressed in different tissues, mostly in intestine, placenta and
CC lung.
CC -!- DEVELOPMENTAL STAGE: Increased expression during development.
CC -!- PTM: Autophosphorylated on threonine and tyrosine residues.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC kinase subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X65198; CAA46318.1; -; mRNA.
CC EMBL; M61177; AAA63486.1; -; mRNA.
CC EMBL; M38194; AAA41123.1; -; mRNA.
CC EMBL; U12008; AAA20009.1; -; mRNA.
CC EMBL; AF155236; AAF71666.1; -; mRNA.
CC PIR; JC1451; JC1451.
CC HSP; P28482; LPW5.
CC SMR; P21708; 30-375.
CC Ensembl; ENSRNOG00000019601; Rattus norvegicus.
CC RGD; 3046; Mapk3.
CC GO; GO:0005829; C:cytosol; IDA.
CC GO; GO:0005626; C:insoluble fraction; IDA.
CC GO; GO:0004707; F:MAP kinase activity; TAS.
CC GO; GO:0000189; P:nuclear translocation of MAPK; IDA.
CC GO; GO:0045941; P:positive regulation of transcription; IE.
CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
CC GO; GO:0007243; P:protein kinase cascade; TAS.
CC InterPro; IPR008349; Erk_1_2_MAPK.
CC InterPro; IPR008351; JNK_MAPK.
CC InterPro; IPR003527; MAP_kin.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_kin_AS.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; Pkinase; 1.
CC PRINTS; PR01770; ERK1ERK2MAPK.
CC PRINTS; PR01772; JNKMAPKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS01351; MAPK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC Alternative splicing; ATP-binding; Cell cycle;
CC Direct protein sequencing; Kinase; Nuclear protein;
KW

```

Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase; Transferase.

DOMAIN 43 331 protein kinase.
NP_BIND 49 57 ATP (By similarity).
ACT_SITE 167 167 Proton acceptor (By similarity).
FT BINDING 72 72 ATP (By similarity).
FT MOD_RES 203 203 Phosphothreonine (activates the kinase) (By similarity).
FT MOD_RES 205 205 phosphotyrosine (activates the kinase) (By similarity).
FT VARSPLIC 340 340 E -> EVSRPPAAGRGISVSPVPVCLCPQ (in isoform B).
CONFLICT 95 95 /FTID=VSP_004830.
SEQUENCE 380 AA; 42981 MW; 53785C4ED722723A CRC64;

Query Match 94.9%; Score 1784; DB 1; Length 380;
Best Local Similarity 91.8%; Pred. No. 2.7e-105;
Matches 347; Conservative 1; Mismatches 10; Indels 20; Gaps 1;

QY 2 AAAAAOAGGGGPRRTGEGVGVGPGVEVMVKGQPFVGPRTYQIQYIGEGAYGMVSSAYD 61
DB 3 AAAAAOAGGGGPRRTGAGVVPVPGVEVMVKGQPFVGPRTYQIQYIGEGAYGMVSSAYD 62
QY 62 HVRKTRVAIKKISPFHQYTCQRTLEIQILLRPHENVIGIRDIRLRASTLEAMRDYIV 121
DB 63 HVRKTRVAIKKISPFHQYTCQRTLEIQILLRPHENVIGIRDIRLRASTLEAMRDYIV 122
QY 122 QDLMETDLYKLLSQQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLK 181
DB 123 QDLMETDLYKLLSQQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLK 182
QY 182 ICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSN 241
DB 183 ICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSN 242
QY 242 RPIFPKGHYLDQLNHILGILGSPSQEDLNCIINMKARNYLSQSPSKTKVAKLFPKSDS 301
DB 243 RPIFPKGHYLDQLNHILGILGSPSQEDLNCIINMKARNYLSQSPSKTKVAKLFPKSDS 302
QY 302 KALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPERLK 341
DB 303 KALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPERLK 362
QY 342 ELIFQETARFQGPVLEAP 359
DB 363 ELIFQETARFQGAPEAP 380

RESULT 4
QAPIY8_RAT PRELIMINARY; PRT; 380 AA.
AC QAPIY8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Extracellular-signal-regulated kinase 1.
GN Name=Mapk3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93013050; PubMed=327976; DOI=10.1016/0378-1119(92)90109-3;
RA Marquardt B., Stabel S.;
RT "Sequence of a rat cDNA encoding the ERK1-MAP kinase."
FL Gene 120:297-299(1992).
DR EMBL; 546779; AAA11604.1; -; mRNA.
KW Kinase.
SQ SEQUENCE 380 AA; 42981 MW; 53785C4ED722723A CRC64;

Query Match 94.9%; Score 1784; DB 2; Length 380;
Best Local Similarity 91.8%; Pred. No. 2.7e-105;
Matches 347; Conservative 1; Mismatches 10; Indels 20; Gaps 1;

QY 2 AAAAAOAGGGGPRRTGEGVGVGPGVEVMVKGQPFVGPRTYQIQYIGEGAYGMVSSAYD 61
DB 3 AAAAAOAGGGGPRRTGAGVVPVPGVEVMVKGQPFVGPRTYQIQYIGEGAYGMVSSAYD 62
QY 62 HVRKTRVAIKKISPFHQYTCQRTLEIQILLRPHENVIGIRDIRLRASTLEAMRDYIV 121
DB 63 HVRKTRVAIKKISPFHQYTCQRTLEIQILLRPHENVIGIRDIRLRASTLEAMRDYIV 122
QY 122 QDLMETDLYKLLSQQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLK 181
DB 123 QDLMETDLYKLLSQQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLK 182
QY 182 ICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSN 241
DB 183 ICDFGLARIADPHDHTGFLTEVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSN 242
QY 242 RPIFPKGHYLDQLNHILGILGSPSQEDLNCIINMKARNYLSQSPSKTKVAKLFPKSDS 301
DB 243 RPIFPKGHYLDQLNHILGILGSPSQEDLNCIINMKARNYLSQSPSKTKVAKLFPKSDS 302
QY 302 KALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPERLK 341
DB 303 KALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPERLK 362
QY 342 ELIFQETARFQGPVLEAP 359
DB 363 ELIFQETARFQGAPEAP 380

RESULT 5
Q7M0H9_CRIGR PRELIMINARY; PRT; 369 AA.
ID Q7M0H9;
AC Q7M0H9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) ERK1 (Fragment).
DE Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Cricetidae; Cricetinae; Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92199340; PubMed=1372523;
RA Meloche S., Pages G., Pouyssegur J.;
RT "Functional expression and growth factor activation of an epitope-tagged p44 mitogen-activated protein kinase, p44(mak).";
RL Mol. Biol. Cell 3:63-71(1992).
CC CC
CC -1- FUNCTION: Kinase involved in a signal transduction pathway that is activated by changes in the osmolarity of the extracellular environment. Plays an essential role in maintaining water homeostasis (By similarity).
CC CC
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR PIR; A60041; A60041.
DR HSSP; Q16539; IOZ1.
DR SMR; Q7M0H9; 19-364.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004685; F:calcium- and calmodulin-dependent protein k. . .; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk 1-2 MAPK.
DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Pro_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.

```
DR pfam: PF00069; Pkinase: 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase: 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_SF; 1.
DR ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 369 369
FT NON_TER 1 1
SQ SEQUENCE 369 AA; 42126 MW; 0BB8FCAAC690B4FF CRC64;

Query Match 94.0%; Score 1767; DB 2; Length 369;
Best Local Similarity 92.4%; Pred. No. 3.2e-104;
Matches 341; Conservative 1; Mismatches 7; Indels 20; Gaps 1;

QY 11 GGEPRTEGVGPGVGEVEMVKGQPPDVGYPTQLQYIGEGAYGMVSSAYDHRVTRVAI 70
Db 1 GGEPRGAAGVPGVGEVEVVGKQPPDVGYPTQLQYIGEGAYGMVSSAYDHRVTRVAI 60
QY 71 KKISPEHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLY 130
Db 61 KKISPEHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLY 120
QY 131 KLLKSQSLNDHCYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDLKITCDFGLARI 190
Db 121 KLLKSQSLNDHCYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDLKITCDFGLARI 180
QY 191 ADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLSNRPIPGKHY 250
Db 181 ADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLSNRPIPGKHY 240
QY 251 LDQLNHILGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSDSKALDLDLRM 310
Db 241 LDQLNHILGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSDSKALDLDLRM 300
QY 311 LTFNPNKRIT-----VAEPTTFAMELDDLPKRLKELIFOETAR 350
Db 301 LTFNPNKRITVEDALAHPLYEQYDPTDFVABEPTDFMELDDLPKRLKELIFOETAR 360
QY 351 FQPGVLEAP 359
Db 361 FQGAPEAP 369

RESULT 6
Q8NHX1 HUMAN
ID Q8NHX1_HUMAN PRELIMINARY; PRT; 357 AA.
AC Q8NHX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-related kinase 1b.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Abersold D.M., Yung Y., Seger R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033607; AAK52329.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q8NHX1; 29-344.
DR Ensembl; ENSG00000102882; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004687; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.

DR InterPro; IPR008351; JNK MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase: 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 357 AA; 40062 MW; 58C92773983ADA79 CRC64;

Query Match 89.6%; Score 1683; DB 2; Length 357;
Best Local Similarity 99.4%; Pred. No. 6.5e-99;
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAAAAAQQGGGEGPRTEGVGPGVGEVEMVKGQPPDVGYPTQLQYIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGEGPRTEGVGPGVGEVEMVKGQPPDVGYPTQLQYIGEGAYGMVSSAY 60
QY 61 DHVTRTRVAIKKISPEHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVTRTRVAIKKISPEHQTYCQRTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
QY 121 VDLMETDLYKLLKSQSLNDHCYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDL 180
Db 121 VDLMETDLYKLLKSQSLNDHCYFLYQILRLGLKYIHSANVLHRLKPSNLLINTTCDL 180
QY 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQLNHILGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSD 300
Db 241 NRPIFGKHYLDQLNHILGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSD 300
QY 301 SKALDLDLRMLTFNPNKRITVAE 323
Db 301 SKALDLDLRMLTFNPNKRITVEE 323

RESULT 7
Q8NHX0 HUMAN
ID Q8NHX0_HUMAN PRELIMINARY; PRT; 335 AA.
AC Q8NHX0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-related kinase 1c.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Abersold D.M., Yung Y., Seger R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033608; AAK52330.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q8NHX0; 29-330.
DR Ensembl; ENSG00000102882; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004687; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
```


DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 335 AA; 38249 MW; A82027A684BA12E1 CRC64;

Query Match 84.8%; Score 1593; DB 2; Length 335;
Best Local Similarity 82.8%; Pred. No. 3.1e-93;
Matches 314; Conservative 0; Mismatches 1; Indels 64; Gaps 2;

QY 1 MAAAAAQQGGGGEPRTEGVGPGVEVMVKQPFDFVGPRTYQIYIGEGAYGMVSSAY 60
DB 1 MAAAAAQQGGGGEPRTEGVGPGVEVMVKQPFDFVGPRTYQIYIGEGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPFHQTYCQRTLRRIQILRRPHENVIGIRDILRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPFHQTYCQRTLRRIQILRRPHENVIGIRDILRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQQLSNDHICFYFLVQILRGLKYIHSANVLHRDLKPSNLLNTTCDL 180
DB 121 VQDLMETDLYKLLKSQQLSNDHICFYFLVQILRGLKYIHSANVLHRDLKPSNLLNTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240

QY 241 NRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPPKSD 300
DB 241 NRPIFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLFPPKSD 300

QY 301 SKALDILLDRMLTFNPNKRIT-----VAERPFTFAMELDDLPKRL 340
DB 259 --ALDILLDRMLTFNPNKRITVEALAHPLYEQYDPTDEPVAESPFTFAMELDDLPKRL 316

QY 341 KELIFQETARFQGVLEAP 359
DB 317 KELIFQETARFQGVLEAP 335

RESULT 8
Q723H5_HUMAN
ID Q723H5_HUMAN PRELIMINARY; PRT; 323 AA.
AC Q723H5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686O0215 (Fragment).
GN Name=DKFZp686O0215;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RA Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Kinase involved in a signal transduction pathway that is
CC activated by changes in the osmolarity of the extracellular
CC environment. Plays an essential role in maintaining water
CC homeostasis (By similarity).
CC EMBL; BX537897; CAD97888.1; -; mRNA.
DR HSP; Q16539; I021.
DR SMR; Q723H5; 2-318.
DR Ensembl; ENSG00000102882; Homo sapiens.

DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK_MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 323 AA; 37501 MW; 53D4EF0913948C1 CRC64;

Query Match 83.1%; Score 1561; DB 2; Length 323;
Best Local Similarity 93.8%; Pred. No. 3.2e-91;
Matches 303; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 57 SSAYDHVRKTRVAIKKISPFHQTYCQRTLRRIQILRRPHENVIGIRDILRASTLEAMR 116
DB 1 SSAYDHVRKTRVAIKKISPFHQTYCQRTLRRIQILRRPHENVIGIRDILRASTLEAMR 60

QY 117 DVYIVQDLMETDLYKLLKSQQLSNDHICFYFLVQILRGLKYIHSANVLHRDLKPSNLLNT 176
DB 61 DVYIVQDLMETDLYKLLKSQQLSNDHICFYFLVQILRGLKYIHSANVLHRDLKPSNLLNT 120

QY 177 TCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILA 236
DB 121 TCDLKICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILA 180

QY 237 EMLSNRPFPFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLF 296
DB 181 EMLSNRPFPFGKHLYDLQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAKLF 240

QY 297 PKSDSKALDILLDRMLTFNPNKRIT-----VAERPFTFAMELDDLP 336
DB 241 PKSDSKALDILLDRMLTFNPNKRITVEALAHPLYEQYDPTDEPVAESPFTFAMELDDLP 300

QY 337 KERLKELIQETARFQGVLEAP 359
DB 301 KERLKELIQETARFQGVLEAP 323

RESULT 9
Q72VK8_BRARE
ID Q72VK8_BRARE PRELIMINARY; PRT; 392 AA.
AC Q72VK8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Mitogen-activated protein kinase 3 (Extracellular signal-regulated
DE kinase 1).
DE Name=mapk3;
GN Brachydanio rerio (Zebrafish) (Danio rerio).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AB; TISSUE=Embryo, and Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=AB; TISSUE=Whole body;
RC NIH MGC Project;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Embryo;
RA Director MGC Project;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
[5]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Tuebingen;
RA Krens S.F.G., Snaar-Jagalska B.E., Spaink H.P.;
RT "Zebrafish ERK1 and ERK2";
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045505; AAH45505.1; -; mRNA.
DR EMBL; BC097073; AAH97073.1; -; mRNA.
DR EMBL; BC066401; AAH66401.1; -; mRNA.
DR EMBL; AY922319; AAY57804.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q7ZVK8; 41-386.
DR ZFIN; ZDB-GENE-040121-1; mapk3.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK_MAPK.
DR InterPro; IPR00719; Prof_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKG; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS0107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 392 AA; 44099 MW; C184681E448E5BBE CRC64;

Query Match 82.6%; Score 1551.5; DB 2; Length 392;
Best Local Similarity 79.9%; Pred. No. 1.6e-90;
Matches 303; Conservative 15; Mismatches 32; Indels 29; Gaps 3;

QY 2 AAAAAOQ-----GGGGERRTTEGVGPGVEVMVKGPFDVGPRTYQIYCEGAY 53
DB 9 AAAGAAGNSSNAGPGGAVAGPGPSGAAGSKPG-LESVKGNFVGPRTYDLYIGEGAY 67
QY 54 GVVSSAYDHRVTRVAIKKISPFHEQTYCQRTLRREIQILLRPHENVIQIRILRASTLE 113
DB 68 GWCVSFAFNVNKRVAIKKISPFHEQTYCQRTLRREIKILLRPHENIIGINDILRAHID 127
QY 114 AMRDYIVQDLMETDLYKLLKSKQQLSNDDHCYFLYQIILGLKLYIHSANVLRDLKPSNLL 173
DB 128 YMRDYIVQDLMETDLYKLLKSKQQLSNDDHCYFLYQIILGLKLYIHSANVLRDLKPSNLL 187
QY 174 INTTCDLKICDPLGARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDISWVGC 233
DB 188 INTTCDLKICDPLGARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKGYTKSIDISWVGC 247
QY 234 ILAEMLSNRPIPPGKHLYDLQNLHILGILGPSQEDLNCIINNKARNYLOSPLSKTKVANA 293
DB 248 ILAEMLSNRPIPPGKHLYDLQNLHILGILGPSQEDLNCIINNKARNYLOSPLSKTKIPWN 307
QY 294 KLFPKSDSKALDLDRLMTFNPKKIT-----VAEEPTFAWELD 333
DB 308 KLFPKADKALDLDRLMTFNPKKINVEQALAHPLYLSQYDPSDEPVAEEPTFNWELD 367
QY 334 DLPKERLKLIFQETARFQ 352
DB 368 DLPKERLKLIFQETARFQ 386
RESULT 10
ID Q9DGR6 BRARE PRELIMINARY; PRT; 391 AA.
AC Q9DGR6;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE ERK1.
GN Name=mapk3; Synonyms=zERK1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hirano T., Fujii R., Hibi M.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB030902; BAB11812.1; -; mRNA.
DR HSSP; P28482; 1PME.
DR SMR; Q9DGR6; 40-385.
DR ZFIN; ZDB-GENE-040121-1; mapk3.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR008351; JNK_MAPK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR PRINTS; PR01772; JNKMAPKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKG; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 391 AA; 44126 MW; 6638F0807B5F9889 CRC64;
Query Match 82.5%; Score 1550.5; DB 2; Length 391;

Best Local Similarity 81.1%; Pred. No. 1.8e-90;
Matches 300; Conservative 17; Mismatches 32; Indels 21; Gaps 2;

QY 3 AAAAGGGGPRRTSGVGPVGEVEMVKGQDFDVGPRYTQYIGEGAGVMSAYDH 62
DB 17 SSAAGGGGAVAGGPGGAAGSKFG-LESVKGQDFDVGPRYTQYIGEGAGVMSAFDN 75

QY 63 VRKTRVAIKKISPFHQYQRTLRREIQILLRPHENVIIGIRDILRASTLEAMRDVYIVQ 122
DB 76 VNKIRVAIKKISPFHQYQRTLRREIKILLRPHENVIIGINDILRAHDVYRDVYIVQ 135

QY 123 DLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKI 182
DB 136 DLMETDLYKLLKTQQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPSNLLINTTCDLKI 195

QY 183 CDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR 242
DB 196 CDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNR 255

QY 243 PIFPGKHLYDLQNLHILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKWAKLFPKSDSK 302
DB 256 PIFPGKHLYDLQNLHILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKWAKLFPKADNK 315

QY 303 ALDLDRLMTFNPVKRIT-----VAEPTTFAMELDDLPKRLKE 342
DB 316 ALDLDRLMTFNPVKRINVRQALAHPLYEQYDPSDEPVAEPTTFNMELDDLPKRLKE 375

QY 343 LIPOETARFQ 352
DB 376 LIPEETARFQ 385

RESULT 11
Q7LZH4_CYPCA
ID Q7LZH4_CYPCA PRELIMINARY; PRT; 392 AA.
AC Q7LZH4;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Extracellular signal-regulated kinase (EC 2.7.-.-) 1 (Extracellular
DE signal regulated protein kinase 1).
GN Name=cERK1;
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=98269030; PubMed=9603989;
RX Hashimoto H., Yokoyama Y., Matsuo Y., Toyohara H., Kohno M.,
RA Sakaguchi M.;
RT "Existence of two isoforms of extracellular signal-regulated kinase in
RT fish."
RL J. Biochem. 123:1031-1035(1998).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB006038; BAD23842.1; -, mRNA.
DR PIR; J00052; J00052.
DR HSSP; Q16539; 1021.
DR SMR; Q7LZH4; 39-386.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004707; F:MAP kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008349; Erk_1_2_MAPK.
DR InterPro; IPR003527; MAP Kin.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF00069; Pkinase; I.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR ProDom; PD000001; Prot kinase; 2.
DR PROSITE; PS01351; MAPK; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 392 AA; 44133 MW; 02E90132E5C347B CRC64;

Query Match 82.5%; Score 1550.5; DB 2; Length 392;
Best Local Similarity 78.5%; Pred. No. 1.8e-90;
Matches 300; Conservative 20; Mismatches 29; Indels 33; Gaps 3;

QY 3 AAAAGGGGPRRTSGVGPV-----PGEVEMVKGQDFDVGPRYTQYIG 50
DB 6 SSAAGGAAGSSSSAAGAGAAVAGGANGAAGPKPG-LESVKGQDFDVGPRYTQYIG 64

QY 51 GAYGVSSAYDVRKTRVAIKKISPFHQYQRTLRREIQILLRPHENVIIGIRDILRAS 110
DB 65 GAYGVCSAFDNVKNIRVAIKKISPFHQYQRTLRREIKILLRPHENVIIGINDILRAS 124

QY 111 TLEAMRDVYIVQDLMETDLYKLLKSQQLSNDHICVFLYQILRLGLKYIHSANVLRDLKPS 170
DB 125 RIEYMRDVYIVQDLMETDLYKLLKTQQLSNDHVCVFLYQILRLGLKYIHSANVLRDLKPS 184

QY 171 NLLINTTCDLKCDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWS 230
DB 185 NLLINTTCDLKCDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWS 244

QY 231 VGCILAEMLSNRPIFGKHLYDLQNLHILGSPSQEDLNCIINMKARNYLQSLPSKTKV 290
DB 245 VGCILAEMLSNRPIFGKHLYDLQNLHILGSPSQEDLNCIINMKARNYLQALPQPKI 304

QY 291 AWAKLFPKSDSKALDLDRLMTFNPVKRIT-----VAEPTTFAM 330
DB 305 PNWKLFPKADNKALDLDRLMTFNPVKRITVEALAHPLYEQYDPSDEPVAEPTTFNM 364

QY 331 ELDDLPKRLKELIPOETARFQ 352
DB 365 ELDDLPKRLKELIPEETARFQ 386

RESULT 12
MK01_BOVIN
ID MK01_BOVIN STANDARD; PRT; 360 AA.
AC P46196;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase
DE 2) (MAPK 2) (MAPK 2) (p42-MAPK) (ERT1).
GN Name=MAPK1; Synonyms=ERK2, PRKM1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=adrenal medulla;
RA Ely C.M., Cox M.E., Her J., Parsons S.J.;
RT "Cloning and sequencing of ERK2 from a bovine adrenal medulla cDNA
RT library."
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Phosphorylates microtubule-associated protein 2 (MAP2).
CC Myelin basic protein (MBP), and Elk-1; may promote entry in the
CC cell cycle.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- ENZYME REGULATION: Activated by phosphorylation on tyrosine and
CC threonine in response to insulin and NGF.
CC -1- PTM: Autophosphorylated on threonine and tyrosine residues, which
CC correlates with a slow and low level of autoactivation (By
CC similarity).

```

CC CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC CC kinase subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: Z14089; CAA78467.1; -; mRNA.
CC PIR: S25011; S25011.
CC HSR: P28482; IPME.
CC SRR: P46196; 16-359.
CC InterPro: IPR008349; Erk_1_2_MAPK.
CC InterPro: IPR003527; MAP_kin.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR008271; Ser_thr_pkin_AS.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; Pkinase_1.
CC PRINTS: PR01770; ERK1ERK2MAPK.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00220; S_TKC; 1.
CC PROSITE: PS01351; MAPK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC ATP-binding; Cell cycle; Kinase; Nucleotide-binding; Phosphorylation;
KW Serine/threonine-protein kinase; Transferase.
FT DOMAIN 25 313 Protein kinase.
FT NP_BIND 31 39 ATP (By similarity).
FT COMPAS 2 9 Poly-Ala.
FT ACT_SITE 149 149 Proton acceptor (By similarity).
FT BINDING 54 54 ATP (By similarity).
FT MOD_RES 185 185 Phosphothreonine (activates the kinase)
FT MOD_RES 187 187 Phosphotyrosine (activates the kinase)
FT MOD_RES 187 187 Phosphotyrosine (By similarity).
SQ SEQUENCE 360 AA; 41376 MW; 58500B2A4E9549DE CRC64;

Query Match 81.9%; Score 1538.5; DB 1; Length 360;
Best Local Similarity 79.1%; Pred. No. 9.7e-90;
Matches 296; Conservative 18; Mismatches 23; Indels 37; Gaps 2;

QY 1 MAAAAAGGGGEPRTGEGVGPVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY 60
DB 1 MAAAAAGAG-----PEMVRGQVDFGVGRYTNLSYIGEGAYGMVCSAY 43

QY 61 DHVKTQVAIKTSPPEHOTYQORTLREIQILLRFRHENVIGIRDIRASTLEAMRDVYI 120
DB 44 DNVNKRVAIKTSPPEHOTYQORTLREIKILLRFRHENVIGINDIRAPTIEQMDVYI 103

QY 121 VQDLMETDLYKLKLSQOLSNHDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTCDL 180
DB 104 VQDLMETDLYKLKLTQHLNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTCDL 163

QY 181 KICDFGLARIADPEHHTGTLTYVATRWYRAPEIMLSNGYTKSDIWSVGCILAEMLS 240
DB 164 KICDFGLARVADPDHDTGTLTYVATRWYRAPEIMLSNGYTKSDIWSVGCILAEMLS 223

QY 241 NRPIFGKHVLDQNLHILGILSPSQEDLNCIIMKARNYLOSLPKTKVAVAKLPKSD 300
DB 224 NRPIFGKHVLDQNLHILGILSPSQEDLNCIIMKARNYLSLPHKNKVPNNRLEFPNAD 283

QY 301 SKALDLDRLMTFNPKNRI-----TVASEPTTFAMELDDLPKRL 340
DB 284 SKALDLDKMLTFNPKHRIEVEQALHPYLEQYDPSDEPVAAPFKFDMLDLPKRL 343

QY 341 KELIFQETARFQPG 354
DB 344 KELIFQETARFQPG 357

```

RESULT 13

```

MK01_HUMAN
ID MK01_HUMAN STANDARD; PRT; 360 AA.
AC P28482;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular
DE signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase
DE 2) (MAP kinase 2) (MAPK 2) (P42-MAPK) (BRT1).
GN Name=MAPK1; Synonyms=ERK2, PRKM1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OC NCBI_TaxID=9606;
RN [1]
RX NUCLEOTIDE SEQUENCE.
RX MEDLINE=92171961; PubMed=1540184;
RA Owaki H., Makar R., Boulton T.G., Cobb M.H., Geppert T.D.;
RT "Extracellular signal-regulated kinases in T cells: characterization
RT of human ERK1 and ERK2 cDNAs."
RL Biochem. Biophys. Res. Commun. 182:1416-1422(1992).
RN [2]
RX NUCLEOTIDE SEQUENCE.
RX MEDLINE=92316223; PubMed=1319925; DOI=10.1016/0014-5793(92)80612-K;
RA Gonzalez F.A., Raden D.L., Rigby M.R., Davis R.J.;
RT "Heterogeneous expression of four MAP kinase isoforms in human
RT tissues."
RL FEBS Lett. 304:170-178(1992).
RN [3]
RX NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RX INTERACTION WITH HIV-1 NEF.
RX PubMed=8794306;
RA Greenway A.L., Aaad A., Mills J., McPhee D.A.;
RT "Human immunodeficiency virus type 1 Nef binds directly to LCK and
RT Mitogen-activated protein kinase, inhibiting kinase activity."
RL J. Virol. 70:6701-6708(1996).
CC -!- FUNCTION: Phosphorylates microtubule-associated protein 2 (MAP2).
CC Myelin basic protein (MBP), and Elk-1; may promote entry in the
CC cell cycle.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium (By similarity).
CC -!- ENZYME REGULATION: Activated by phosphorylation on tyrosine and
CC threonine in response to insulin and NGF.
CC -!- SUBUNIT: Binds to HIV-1 Nef through its SH3 domain. This
CC interaction inhibits its tyrosine-kinase activity.
CC -!- PTM: Autophosphorylated on threonine and tyrosine residues, which
CC correlates with a slow and low level of autoactivation (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP

```

kinase subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; M84489; AAA58459.1; -; mRNA.
DR EMBL; Z11694; CAA7752.1; -; mRNA.
DR EMBL; Z11695; CAA7753.1; ALT_INIT; mRNA.
DR EMBL; BC017832; AAL17832.1; -; mRNA.
DR PIR; JQ1400; JQ1400.
DR PDB; 1PME; X-ray; @=1-360.
DR OGP; P28482; -.
DR Ensembl; ENSG00000100030; Homo sapiens.
DR HGNC; HGNC:6871; MAPK1.
DR H-InvDB; HIX0016281; -.
DR Reactome; P28482; -.
DR MIM; 176948; -.
DR GO; GO:0004707; F:MAP kinase activity; TAS.
DR GO; GO:0006935; P:chemotaxis; TAS.
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR GO; GO:0006950; P:response to stress; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR GO; GO:0007268; P:synaptic transmission; TAS.
DR InterPro; IPR008349; Erk 1,2 MAPK.
DR InterPro; IPR003527; MAP_kin_
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Chtr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR01770; ERK1ERK2MAPK.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW 3D-structure; ATP-binding; Cell Cycle; Kinase; Nucleotide-binding;
KW Phosphorylation; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 25 313
FT NP_BIND 31 39
FT COMPBIAS 2 9
FT ACT_SITE 149 149
FT BINDING 54 54
FT MOD_RES 185 185
FT MOD_RES 187 187
FT CONFLICT 91 91
FT TURN 23 24
FT STRAND 25 30
FT STRAND 39 44
FT TURN 45 47
FT STRAND 50 56
FT TURN 59 60
FT TURN 62 77
FT TURN 81 82
FT STRAND 83 83
FT STRAND 88 90
FT TURN 95 97
FT STRAND 101 106
FT STRAND 110 111
FT HELIX 112 118
FT TURN 123 142
FT TURN 143 144
FT STRAND 145 146
FT STRAND 152 154
FT STRAND 155 157
FT TURN 159 160
FT STRAND 163 165
FT TURN 169 170

FT STRAND 173 173
FT HELIX 176 176
FT STRAND 180 180
FT TURN 183 184
FT HELIX 191 193
FT HELIX 196 198
FT TURN 199 201
FT STRAND 202 202
FT TURN 207 207
FT HELIX 208 223
FT HELIX 233 244
FT HELIX 249 253
FT TURN 254 254
FT HELIX 258 265
FT TURN 266 266
FT HELIX 275 278
FT TURN 280 281
FT HELIX 284 293
FT TURN 298 300
FT HELIX 304 308
FT TURN 309 309
FT HELIX 311 313
FT TURN 314 316
FT HELIX 319 321
FT HELIX 340 350
FT HELIX 352 354
FT TURN 356 358
SQ SEQUENCE 360 AA; 41390 MW; E85D0B2A5D2D724E CRC64;
Query Match 81.9%; Score 1538.5; DB 1; Length 360;
Best Local Similarity 79.1%; Pred. No. 9.7e-90;
Matches 296; Conservative 18; Mismatches 23; Indels 37; Gaps 2;
Qy 1 MAAAAAQQGGGGPRRTEGVGPGVEVMVKQPPDVGPRTYQLQVIGSGAYGMVSSAY 60
Db 1 MAAAAAAGAG-----PENVRQVFDVGPRTNLSYIGSGAYGMVCSAY 43
Qy 61 DHVRKTRVAIKKISPFEPHQYQCTRLREIQILIRFRHENVIGRIDILRASTLEAMRDVYI 120
Db 44 DNVKVRVAIKKISPFEPHQYQCTRLREIKILLRFRHENIIGINDIIRAPTIEQMKDVI 103
Qy 121 VQDLMETDLYKLKSKQSLNDHICYFLYQLRLGKVIHNSANVLRDLKPSNLLINTTCDL 180
Db 104 VQDLMETDLYKLKLTQHLSDNDHICYFLYQLRLGKVIHNSANVLRDLKPSNLLNTTCDL 163
Qy 181 KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 240
Db 164 KICDFGLARVADPDHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS 223
Qy 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSIPSTKVAKLFPKSD 300
Db 224 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINLKARNYLLSLPHKKNVPMNRLFPNAD 283
Qy 301 SKALDLDRLMTENPKRITV-----AEPPTFAMELDDLPKRL 340
Db 284 SKALDLDRLMTFNPHKRIEVEQALAHPLYEQYDPSDEPIAEAPFKFDMELDDLPKRL 343
Qy 341 KELIFQETARFQPG 354
Db 344 KELIFQETARFQPG 357
RESULT 14
Q8UWG6 CHICK
ID Q8UWG6 CHICK PRELIMINARY; PRT; 368 AA.
AC Q8UWG6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular signal-regulated kinase 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC	Gallus.	DT	13-SEP-2005 (Rel. 48, Last annotation update)
OX	NCBI_TaxID=9031;	DE	Mitogen-activated protein kinase 1 (EC 2.7.1.37) (Extracellular
RN	[1]	DE	signal-regulated kinase 2) (ERK-2) (Mitogen-activated protein kinase
RP	NUCLEOTIDE SEQUENCE.	DE	2) (MAP kinase 2) (MAPK 2) (p42-MAPK) (ERT1).
RC	TISSUE=Pineal gland;	OS	Name=Mapki; Synonyms=Erk2, Mapk, Prkml;
RA	Kasahara T., Higashi C., Okano T., Fukada Y.;	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
RL	Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.	OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
CC	1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.	OC	Muroidea; Muridae; Murinae; Mus.
CC	1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.	OX	NCBI_TaxID=10090;
DR	EMBL; AY033635; AAK56503.1; -; mRNA.	RN	[1]
DR	HSSP; P28482; 1PME.	RP	NUCLEOTIDE SEQUENCE.
DR	SMR; Q8UWG6; 24-367.	RC	STRAIN=Swiss; TISSUE=Fibroblast;
DR	Ensemble; ENSGALG00000001501; Gallus gallus.	RX	MEDLINE=91305126; PubMed=1649458;
DR	GO; GO:003524; F:ATP binding; IEA.	RA	Her J.-H., Wu J.-S., Rall T.B., Sturgill T.W., Weber M.J.;
DR	GO; GO:0004707; F:MAP kinase activity; IEA.	RT	"Sequence of pp42/MAP kinase, a serine/threonine kinase regulated by
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.	RT	tyrosine phosphorylation."
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.	RL	Nucleic Acids Res. 19:3743-3743(1991).
DR	InterPro; IPR008349; Erk 1.2 MAPK.	RN	[2]
DR	InterPro; IPR003527; MAP_kin.	RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
DR	InterPro; IPR000719; Prot_kinase.	RC	STRAIN=C57BL/6J, and NOD; TISSUE=Head, Thymus, and Urinary bladder;
DR	InterPro; IPR008271; Ser_Thr_kinase.	RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
DR	InterPro; IPR002290; Ser_thr_kinase.	RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
DR	Pfam; PF00069; Pkinase; 1.	RA	Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
DR	PRINTS; PR01770; ERK1ERK2MAPK.	RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
DR	ProDom; PD000001; Prot_kinase; 1.	RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
DR	SMART; SM00220; S_TKc_1.	RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
DR	PROSITE; PS01351; MAPK; 1.	RA	Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
DR	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.	RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
DR	PROSITE; PS01008; PROTEIN_KINASE_ST; 1.	RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
KW	ATP-binding; Kinase; Nucleotide-binding;	RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
KW	Serine/threonine-protein kinase; Transferase.	RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
SQ	SEQUENCE 368 AA; 41942 MW; A6F3A3D27C6C9A44 CRC64;	RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
		RA	Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
		RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
		RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
		RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
		RA	Sulcanea R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
		RA	Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
		RA	Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
		RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
		RA	Hirozaki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
		RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
		RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
		RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
		RA	Birney E., Hayashizaki Y.;
		RT	"Analysis of the mouse transcriptome based on functional annotation of
		RT	60,770 full-length cDNAs."
		RL	Nature 420:563-573(2002).
		RN	[3]
		RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
		RC	TISSUE=Eye;
		RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
		RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
		RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
		RA	Altschul S.F., Aebersold B., Buetow K.H., Schaefer C.F., Bhat N.K.,
		RA	Hopkins R.F., Zeeberg H., Moore T., Max S.I., Wang J., Hsieh F.,
		RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
		RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
		RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
		RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
		RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
		RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
		RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
		RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
		RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
		RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
		RA	Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
		RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.,
		RT	"Generation and initial analysis of more than 15,000 full-length human

RESULT 15

MOUSE
ID MK01_MOUSE
AC P63085; P27703;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)

DR	PROSITE; PS60108; PROTEIN KINASE_S7; I.
KW	ATP-binding; Cell cycle; Direct protein sequencing; Kinase;
KW	Nucleotide-binding; Phosphorylation; Serine/threonine-protein kinase;
KW	Transferase.
FT	DOMAIN 23 311 Protein kinase.
FT	NP_BIND 29 37 ATP (By similarity).
FT	COMPAS 2 7 Poly-Ala.
FT	ACT_SITE 147 147 By similarity.
FT	BINDING 52 52 ATP (By similarity).
FT	MOD_RES 183 183 Phosphothreonine (activates the kinase).
FT	MOD_RES 185 185 Phosphotyrosine (activates the kinase).
FT	SEQUENCE 358 AA; 41276 MW; 3BBCF2247IEDBA0B CRC64;
QY	Query Match 81.3%; Score 1528.5; DB 1; Length 358;
DB	Best Local Similarity 78.6%; Pred. No. 4.2e-89;
DB	Matches 294; Conservative 19; Mismatches 22; Indels 39; Gaps 2;
QY	1 MAAAAQQGGGGEPRRTGEGVPGVEVMVGKQPFDVGPRVTQLQYIGEGAGMVSSAY 60 : :
DB	1 MAAAAAAG-----PEWVRGVDFDVGPRVTNLSYIGEGAGMVC SAY 41 : :
QY	61 DHVRKTRVAIKKISPFPHQTQCORTLREIQILLRFRHENVIIGIRDILRASTLEAMRDVYI 120 ::: : :::
DB	42 DNLNKKRVAAIKKISPFPHQTQCORTLREIKILLRFRHENIIIGINDIRAPTIIEQM KD VYI 101 ::: : :::
QY	121 VDQLMETDLYKLLKSQOLSNDHI CYFLYQILAGLK YIHSA NVLRDLKPSNLLINTTC DL 180 ::: : :::
DB	102 VDQLMETDLYKLLKTQHLSNDHI CYFLYQILAGLK YIHSA NVLRDLKPSNLLINTTC DL 161 ::: : :::
QY	181 KICDFGLARIADPEHDHTGFLT EYVATRWYRAPEIM LNSKG YTKSIDIVS VGCILAEMLS 240 ::: : :::
DB	162 KICDFGLARVADPDHDTGFLT EYVATRWYRAPEIM LNSKG YTKSIDIVS VGCILAEMLS 221 ::: : :::
QY	241 NRPIFPKGHYLDQLNHILGILGPSQSDLCNIIMKARNYLQSLPKSKTVAKWLPEKSD 300 ::: : :::
DB	222 NRPIFPKGHYLDQLNHILGILGPSQSDLCNIIMKARNYLQSLPKSKTVAKWLPEKSD 281 ::: : :::
QY	301 SKALDLLDRMLTFNPNKRITV-----AEEPTFFAMELDDLPKERL 340 ::: : :::
DB	282 SKALDLLDKMLTFNPKRIEVEQAHPHYEQYYDPSPDEPIAEAPKFDMELDDLPEKL 341 ::: : :::
QY	341 KELIFQETARPOFG 354 :
DB	342 KELIFETARFOFG 355 :

Search completed: February 6, 2006, 15:34:39
Job time : 172.223 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:34:56 ; Search time 36.2104 Seconds
(without alignments)
819.670 Million cell updates/sec

Title: US-10-623-108-4

Perfect score: 1879

Sequence: 1 MAAAAAGGGGGPRTEGV.....LKELIFQETARFQGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5 COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/6 COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/H COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/RE COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1859	98.9	379	2	US-08-622-277A-8
2	1859	98.9	379	2	US-09-411-628-14
3	1859	98.9	379	2	US-10-174-794-14
4	1859	98.9	379	2	US-09-538-092-1021
5	1853	98.6	379	2	US-09-025-580-25
6	1853	98.6	379	2	US-09-642-749-25
7	1853	98.6	631	2	US-09-417-197-39
8	1787	95.1	624	2	US-09-417-197-57
9	1735	92.3	367	1	US-08-176-620A-11
10	1735	92.3	367	1	US-08-461-985-11
11	1734	92.3	355	2	US-09-459-953A-8
12	1734	92.3	355	2	US-09-393-212-8
13	1723	91.7	365	1	US-08-176-620A-2
14	1723	91.7	365	1	US-08-463-862-2
15	1723	91.7	365	1	US-08-461-985-2
16	1723	91.7	365	1	US-08-458-887-2
17	1723	91.7	365	2	US-08-932-787B-2
18	1723	91.7	365	2	US-08-932-012C-2
19	1723	91.7	365	2	US-08-888-818C-2
20	1553.5	82.7	380	2	US-09-949-016-9919
21	1548.5	82.4	380	2	US-08-459-953A-9
22	1548.5	82.4	380	2	US-09-393-212-9
23	1538.5	81.9	360	2	US-09-457-040B-3
24	1538.5	81.9	360	2	US-09-538-092-1028
25	1530	81.4	415	1	US-08-176-620A-4
26	1530	81.4	415	1	US-08-463-862-4
27	1530	81.4	415	1	US-08-461-985-4

28	1530	81.4	415	1	US-08-458-887-4
29	1530	81.4	415	2	US-08-932-787B-4
30	1530	81.4	415	2	US-08-932-012C-4
31	1530	81.4	415	2	US-08-888-818C-4
32	1529.5	81.4	360	2	US-08-622-277A-14
33	1528.5	81.3	358	1	US-08-176-620A-12
34	1528.5	81.3	358	1	US-08-461-985-12
35	1528.5	81.3	358	2	US-09-457-040B-4
36	1528.5	81.3	364	2	US-09-457-197-59
37	1528.5	81.3	604	2	US-09-417-197-41
38	1528.5	81.3	605	2	US-09-457-040B-35
39	1515.5	80.7	365	2	US-09-025-580-2
40	1509.5	80.3	360	2	US-09-642-749-2
41	1509.5	80.3	360	2	US-08-176-620A-8
42	905	48.2	183	1	US-08-461-985-8
43	905	48.2	183	2	US-08-932-787B-16
44	905	48.2	183	2	US-08-932-012C-16
45	905	48.2	183	2	US-08-932-012C-16

ALIGNMENTS

RESULT 1
US-08-622-277A-8
; Sequence 8, Application US/08622277A
; Patent No. 6001580
; GENERAL INFORMATION:
; APPLICANT: Tani, Akiyoshi
; APPLICANT: Ichimori, Yuzo
; TITLE OF INVENTION: Method For Assaying MAP Kinase
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/622,277A
; FILING DATE: 27-MAR-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 320577-1995
; FILING DATE: 08-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 305456-1995
; FILING DATE: 24-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 070125-1995
; FILING DATE: 28-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Toohy, Kimberlin M.
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 04221.0039-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
US-08-622-277A-8

Query Match 98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRTEGVGPGVEVMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGVGPGVEVMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240

QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAKLPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAKLPKSD 300

QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPEPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNKRIT-----VAEPEPTFAMELDDLPKRL 340

QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 2
US-09-411-628-14
; Sequence 14, Application US/09411628
; Patent No. 6428994
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-411-628-14

Query Match 98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRTEGVGPGVEVMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGVGPGVEVMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240

DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAKLPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAKLPKSD 300

QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPEPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNKRITVEEALAHPLYEQYDPTDEPVAEPTFAMELDDLPKRL 360

QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 3
US-10-174-794-14
; Sequence 14, Application US/10174794
; Patent No. 6664086
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; TITLE OF INVENTION: CDNA, GENOMIC, AND PREDICTED PROTEIN
; FILE REFERENCE: 13761-707
; CURRENT APPLICATION NUMBER: US/10/174,794
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/411,628
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 60/102,906
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-174-794-14

Query Match 98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRTEGVGPGVEVMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60
DB 1 MAAAAQGGGGPRTEGVGPGVEVMVKQPPDVGPRTYQIQYIGEGAYGMVSSAY 60

QY 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
DB 61 DHVRKTRVAIKKISPEHOTYQORTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120

QY 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
DB 121 VQDLMETDLYKLLKSQOLSNNDHCYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
DB 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240

QY 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAKLPKSD 300
DB 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYQSLPSKTKVAKLPKSD 300

QY 301 SKALDLDRLMTFNPKNKRIT-----VAEPEPTFAMELDDLPKRL 340
DB 301 SKALDLDRLMTFNPKNKRITVEEALAHPLYEQYDPTDEPVAEPTFAMELDDLPKRL 360

QY 341 KELIFOETARFQGVLEAP 359
DB 361 KELIFOETARFQGVLEAP 379

RESULT 4
US-09-538-092-1021

```
; Sequence 1021, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormat Version 0.9
; SEQ ID NO 1021
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number P27361
; US-09-538-092-1021

Query Match          98.9%; Score 1859; DB 2; Length 379;
Best Local Similarity 94.7%; Pred. No. 1.2e-185;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

Qy 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKTSPFEHQYTCQRTLEIQILRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKTSPFEHQYTCQRTLEIQILRFRHENVIGIRDILRASTLEAMRDVYI 120
Qy 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Qy 241 NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Db 241 NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFQETARFQPGVLEAP 359
Db 361 KELIFQETARFQPGVLEAP 379

RESULT 5
US-09-025-580-25
; Sequence 25, Application US/09025580
; Patent No. 6162613
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; APPLICANT: Fox, Ted
; APPLICANT: Wilson, Keith Phillip
; APPLICANT: Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; TITLE OF INVENTION: Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
```

```
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-580-25

Query Match          98.6%; Score 1853; DB 2; Length 379;
Best Local Similarity 94.5%; Pred. No. 5.2e-185;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

Qy 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
Db 1 MAAAAAGGGGGPRRTGEGVGPGEVEMVKGPFDVGPRTQLOVIGSGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKTSPFEHQYTCQRTLEIQILRFRHENVIGIRDILRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKTSPFEHQYTCQRTLEIQILRFRHENVIGIRDILRASTLEAMRDVYI 120
Qy 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VDLMETDLYKLKLSQQLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEVATRWYRAPEIMLNSKGYTKSIDWSVGCILAEMLS 240
Qy 241 NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Db 241 NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAKLFPKSD 300
Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFQETARFQPGVLEAP 359
Db 361 KELIFQETARFQPGVLEAP 379

RESULT 6
US-09-642-749-25
; Sequence 25, Application US/09642749
; Patent No. 6849716
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; APPLICANT: Fox, Ted
; APPLICANT: Wilson, Keith Phillip
; APPLICANT: Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; TITLE OF INVENTION: Serine/Threonine Kinases and Tyrosine Kinase
```

; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/642,749
; APPLICATION NUMBER: US/09/642,749
; FILING DATE: 18-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-642-749-25

Query Match 98.6%; Score 1853; DB 2; Length 379;
Best Local Similarity 94.5%; Pred. No. 5.2e-185;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRRTGPGVPGVEVMVKGPDPVGPRTYQLQYIGEGAYGMVSSAY 60
Db 1 MAAAAQGGGGPRRTGPGVPGVEVMVKGPDPVGPRTYQLQYIGEGAYGMVSSAY 60
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRPHENVIGIRDIRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRPHENVIGIRDIRASTLEAMRDVYI 120
QY 121 VQDLMETDLYKLLKSKQSLNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSKQSLNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYQLSLPSKTKVAKLPFKSD 300
Db 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYQLSLPSKTKVAKLPFKSD 300
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKERL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDTPDPEVAEPPFTFAMELDDLPKERL 360
QY 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 7
US-09-417-197-39
; Sequence 39, Application US/09417197
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mERK1-F64L-S65T-GFP fusion
US-09-417-197-57
Query Match 95.1%; Score 1787; DB 2; Length 624;

; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Erk1 fusion construct
US-09-417-197-39

Query Match 98.6%; Score 1853; DB 2; Length 631;
Best Local Similarity 94.5%; Pred. No. 1.1e-184;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MAAAAQGGGGPRRTGPGVPGVEVMVKGPDPVGPRTYQLQYIGEGAYGMVSSAY 60
Db 253 MAAAAQGGGGPRRTGPGVPGVEVMVKGPDPVGPRTYQLQYIGEGAYGMVSSAY 312
QY 61 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRPHENVIGIRDIRASTLEAMRDVYI 120
Db 313 DHVRKTRVAIKKISPEHQTYCQRTLRREIQILLRPHENVIGIRDIRASTLEAMRDVYI 372
QY 121 VQDLMETDLYKLLKSKQSLNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 180
Db 373 VQDLMETDLYKLLKSKQSLNDHCYFLYQILRLGKYIHSANVLHRDLKPSNLLINTTCDL 432
QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240
Db 433 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 492
QY 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYQLSLPSKTKVAKLPFKSD 300
Db 493 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYQLSLPSKTKVAKLPFKSD 552
QY 301 SKALDLDRLMTFNPKNRIT-----VAEPPFTFAMELDDLPKERL 340
Db 553 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYDTPDPEVAEPPFTFAMELDDLPKERL 612
QY 341 KELIFQETARFQGVLEAP 359
Db 613 KELIFQETARFQGVLEAP 631

RESULT 8
US-09-417-197-57
; Sequence 57, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mERK1-F64L-S65T-GFP fusion
US-09-417-197-57
Query Match 95.1%; Score 1787; DB 2; Length 624;

```
Best Local Similarity 91.8%; Pred. No. 8.9e-178;
Matches 347; Conservative 1; Mismatches 10; Indels 20; Gaps 1;

QY 2 AAAAAAGGGGGRRRTGGVPGVGVEMVKGQPFVDPVRYTQLOVIGEGAYGMVSSAYD 61
Db 3 AAAAAAGGGGGRRRTAGVVPVPGVGVEMVKGQPFVDPVRYTQLOVIGEGAYGMVSSAYD 62
QY 62 HVKTRVAIKKISPFHEQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIV 121
Db 63 HVKTRVAIKKISPFHEQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIV 122
QY 122 QDLMETDLYKLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLK 181
Db 123 QDLMETDLYKLLKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLK 182
QY 182 ICDPGLARIADPDHDTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 241
Db 183 ICDPGLARIADPDHDTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSN 242
QY 242 RPIFPKGHYLDQNLHILGILGSPSQEDLNCIIINMKARNYLQSLPSKTKVAVAKLPPKSDS 301
Db 243 RPIFPKGHYLDQNLHILGILGSPSQEDLNCIIINMKARNYLQSLPSKTKVAVAKLPPKSDS 302
QY 302 KALDLDRLMTFNPKNRIT-----VAEPPTFAMELDLDPKRLK 341
Db 303 KALDLDRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFDMELDLDLPKRLK 362
QY 342 ELIFQETARFQPGVLEAP 359
Db 363 ELIFQETARFQGAPEGP 380

RESULT 9
US-08-176-620A-11
; Sequence 11, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,620A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-176-620A-11
Query Match 92.3%; Score 1735; DB 1; Length 367;
Best Local Similarity 91.8%; Pred. No. 1.1e-172;
Matches 337; Conservative 1; Mismatches 9; Indels 20; Gaps 1;

QY 13 EPRRTGEGVPGVGVEMVKGQPFVDPVRYTQLOVIGEGAYGMVSSAYDHRKTRVAIKK 72
Db 1 EPRGTAGVVPVPGVGVEMVKGQPFVDPVRYTQLOVIGEGAYGMVSSAYDHRKTRVAIKK 60
QY 73 ISPFHEQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYK 132
Db 61 ISPFHEQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYK 120
QY 133 LKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIAD 192
Db 121 LKSQLSNDHICVFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIAD 180
QY 193 PEHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYD 252
Db 181 PEHDHTGFLTEVATRWYRAPEIMLSNGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYD 240
QY 253 QLNHLILGILGSPSQEDLNCIIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLDRLMT 312
Db 241 QLNHLILGILGSPSQEDLNCIIINMKARNYLQSLPSKTKVAVAKLFPKSDSKALDLDRLMT 300
QY 313 FNPKNRIT-----VAEPPTFAMELDLDPKRLKELIFQETARFQ 352
Db 301 FNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFDMELDLDLPKRLKELIFQETARFQ 360
QY 353 PGVLEAP 359
Db 361 PGAPAP 367

RESULT 10
US-08-461-985-11
; Sequence 11, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Panayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,985
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,620
; FILING DATE: 03-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6526-123
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-461-985-11

Query Match          92.3%; Score 1735; DB 1; Length 367;
Best Local Similarity 91.8%; Pred. No. 1.1e-172; Indels 20; Gaps 1;
Matches 337; Conservative 1; Mismatches 9;

QY 13 EPRTEGVGPGVGEVEMVKGPFDVGPRTYQLQYIGEGAYGMVSSAYDHVRKTRVAIKK 72
DB 1 EPRGTAGVVPVPGVEVVKGPFDVGPRTYQLQYIGEGAYGMVSSAYDHVRKTRVAIKK 60
QY 73 ISPFHQTYCQRTLRIOILLRFRHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLYKL 132
DB 61 ISPFHQTYCQRTLRIOILLRFRHENVIGIRDLIRAPTLEAMRDVYIVQDLMETDLYKL 120
QY 133 LKSQLSNDHICVFLYQILRGLKVIHSANVLHRLDKPSNLLINTTCDLKI CDFGLARIAD 192
DB 121 LKSQLSNDHICVFLYQILRGLKVIHSANVLHRLDKPSNLLINTTCDLKI CDFGLARIAD 180
QY 193 PEHDHTGFLTYYVATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYD 252
DB 181 PEHDHTGFLTYYVATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYD 240
QY 253 QLNHIILGILGPSQEDLNCIINMKARNYLSQPSKTKVAVAKLFPKSDSKALDLDRLMT 312
DB 241 QLNHIILGILGPSQEDLNCIINMKARNYLSQPSKTKVAVAKLFPKSDSKALDLDRLMT 300
QY 313 FPNKRIT-----VAEPPFTFAMELDLDPKERLKELIFOETARFQVLEAP 352
DB 301 FPNKRITVEEALAHPLYEQYDPTDPEVAEPPFTFAMELDLDPKERLKELIFOETARFQ 360
QY 353 PGVLEAP 359
DB 361 PGAEAP 367

RESULT 11
US-08-459-953A-8
; Sequence 8, Application US/08459953A
; Patent No. 6030822
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/459,953A
; FILING DATE: June 2, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,494
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 209/267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-953A-8

Query Match          92.3%; Score 1734; DB 2; Length 355;
Best Local Similarity 94.4%; Pred. No. 1.3e-172;
Matches 335; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 25 PGEVEMVKGPFDVGPRTYQLQYIGEGAYGMVSSAYDHVRKTRVAIKKISPEHQTYCOR 84
DB 1 PGEVEMVKGPFDVGPRTYQLQYIGEGAYGMVSSAYDHVRKTRVAIKKISPEHQTYCOR 60
QY 85 TUREIQILRFRHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC 144
DB 61 TUREIQILRFRHENVIGIRDLIRASTLEAMRDVYIVQDLMETDLYKLLKSQOLSNDHIC 120
QY 145 YELYQILRGLKYIHSANVLHRLDKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLT 204
DB 121 YELYQILRGLKYIHSANVLHRLDKPSNLLINTTCDLKI CDFGLARIADPEHDHTGFLT 180
QY 205 VATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYDQNLHILGILGSP 264
DB 181 VATRYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYDQNLHILGILGSP 240
QY 265 SQEDLNCIINMKARNYLSQPSKTKVAVAKLFPKSDSKALDLDRLMTFNPKNRIT- 320
DB 241 SQEDLNCIINMKARNYLSQPSKTKVAVAKLFPKSDSKALDLDRLMTFNPKNRITVEEA 300
QY 321 -----VAEPPFTFAMELDLDPKERLKELIFOETARFQVLEAP 359
DB 301 LAHPYLEQYDPTDPEVAEPPFTFAMELDLDPKERLKELIFOETARFQVLEAP 355

RESULT 12
US-09-393-212-8
; Sequence 8, Application US/09393212
; Patent No. 6579972
; GENERAL INFORMATION:
; APPLICANT: Lechner, Cornelia
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: EXTRACELLULAR SIGNAL-RELATED
; TITLE OF INVENTION: KINASE, SEQUENCES, AND
; TITLE OF INVENTION: METHODS OF PRODUCTION
; AND USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/393,212
FILING DATE: 09-Sep-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/459,953A
FILING DATE: June 2, 1995
APPLICATION NUMBER: 08/029,494
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/267
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-393-212-8

Query Match 92.3%; Score 1734; DB 2; Length 355;
Best Local Similarity 94.4%; Pred. No. 1.3e-172;
Matches 335; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
QY 25 PGEVMVGKQPFVGVPRYTQLYIGEGAYGMVSSAYDHRVTRVAIKKISPFHEHTYQCR 84
DB 1 PGEVMVGKQPFVGVPRYTQLYIGEGAYGMVSSAYDHRVTRVAIKKISPFHEHTYQCR 60
QY 85 TLREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLSQSLNDHIC 144
DB 61 TLREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLSQSLNDHIC 120
QY 145 YFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLTEY 204
DB 121 YFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPEHDHTGFLTEY 180
QY 205 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQLNHLILGILGSP 264
DB 181 VATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQLNHLILGILGSP 240
QY 265 SQEDLNCIINNKARNYIQLSPSKTKVAVAKLFPKSDSKALDLDLDRMLTENPNKRIT--- 320
DB 241 SQEDLNCIINNKARNYIQLSPSKTKVAVAKLFPKSDSKALDLDLDRMLTENPNKRITVERA 300
QY 321 -----VAEPPFTFAMELDLDPKRLKELIFQETARFPQGVLEAP 359
DB 301 LAHPYLEQYDPTDEPVAEPPFTFAMELDLDPKRLKELIFQETARFPQGVLEAP 355

RESULT 13
US-08-176-620A-2
; Sequence 2, Application US/08176620A
; Patent No. 5595904
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.
; APPLICANT: Cobb, Melanie H.
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Nye, Steven
; APPLICANT: Fanayotatos, Nikos
; TITLE OF INVENTION: A Family of Map2 Protein Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,620A
FILING DATE: 03-JAN-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Mierock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-620A-2

Query Match 91.7%; Score 1723; DB 1; Length 365;
Best Local Similarity 91.8%; Pred. No. 1.9e-171;
Matches 335; Conservative 1; Mismatches 9; Indels 20; Gaps 1;
QY 15 RRTGEGVPGVGEVVMVGKQPFVGVPRYTQLYIGEGAYGMVSSAYDHRVTRVAIKKIS 74
DB 1 RGTAGVVPVPGVGEVVMVGKQPFVGVPRYTQLYIGEGAYGMVSSAYDHRVTRVAIKKIS 60
QY 75 PPEHTQYVCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLK 134
DB 61 PPEHTQYVCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYIVQDLMETDLYKLLK 120
QY 135 SQSLNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPE 194
DB 121 SQSLNDHICYFLYQILRLGLKYIHSANVLHRDLKPSNLLINTTCDLKICDFGLARIADPE 180
QY 195 HDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQL 254
DB 181 HDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHYYLDQL 240
QY 255 NHLIGILGSPSQEDLNCIINNKARNYIQLSPSKTKVAVAKLFPKSDSKALDLDLDRMLTFN 314
DB 241 NHLIGILGSPSQEDLNCIINNKARNYIQLSPSKTKVAVAKLFPKSDSKALDLDLDRMLTFN 300
QY 315 PNKRIT-----VAEPPFTFAMELDLDPKRLKELIFQETARFPQGVLEAP 354
DB 301 PNKRITVERAHPYLEQYDPTDEPVAEPPFTFAMELDLDPKRLKELIFQETARFPQGV 360
QY 355 VLEAP 359
DB 361 APEAP 365

RESULT 14
US-08-463-862-2
; Sequence 2, Application US/08463862
; Patent No. 5776751
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G. et al.
; TITLE OF INVENTION: A FAMILY OF MAP2 PROTEIN KINASES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,862
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,544
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-049
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-862-2

Query Match 91.7%; Score 1723; DB 1; Length 365;
Best Local Similarity 91.8%; Pred. No. 1.9e-171;
Matches 335; Conservative 1; Mismatches 9; Indels 20; Gaps 1;

QY	15	RRTGVPVPGVEVWVGQPPDVGPRYTQLOYIGEGAYGMVSSAYDHYRKRTRVAIKKIS	74
Db	1	RGTAGVVPVPGVEVWVGQPPDVGPRYTQLOYIGEGAYGMVSSAYDHYRKRTRVAIKKIS	60
QY	75	PFEHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK	134
Db	61	PFEHQYTCQRTLEIQILLRFRHENVIGIRDILRAPLEAMRDVYIVQDLMETDLYKLLK	120
QY	135	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	194
Db	121	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	180
QY	195	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRRPIFPKGHYLDOL	254
Db	181	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRRPIFPKGHYLDOL	240
QY	255	NHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVAAWAKLFPKSDSKALDLDRLMTFN	314
Db	241	NHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVAAWAKLFPKSDSKALDLDRLMTFN	300
QY	315	PNKRIT-----VAEPPTFAMELDLDPKRLKELIFQETARFQPG	354
Db	301	PNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFDMELDDL PKRLKELIFQETARFQPG	360
QY	355	VLEAP 359	
Db	361	APEAP 365	

RESULT 15

US-08-461-985-2
; Sequence 2, Application US/08461985
; Patent No. 5872006
; GENERAL INFORMATION:
; APPLICANT: Boulton, Teri G.

APPLICANT: Cobb, Melanie H.
APPLICANT: Yancopoulos, George D.
APPLICANT: Nye, Steven
APPLICANT: Panayotatos, Nikos
TITLE OF INVENTION: A Family of Map2 Protein Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,985
FILING DATE: 05-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,620
FILING DATE: 03-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6526-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-985-2

Query Match 91.7%; Score 1723; DB 1; Length 365;
Best Local Similarity 91.8%; Pred. No. 1.9e-171;
Matches 335; Conservative 1; Mismatches 9; Indels 20; Gaps 1;

QY	15	RRTGVPVPGVEVWVGQPPDVGPRYTQLOYIGEGAYGMVSSAYDHYRKRTRVAIKKIS	74
Db	1	RGTAGVVPVPGVEVWVGQPPDVGPRYTQLOYIGEGAYGMVSSAYDHYRKRTRVAIKKIS	60
QY	75	PFEHQYTCQRTLEIQILLRFRHENVIGIRDILRASTLEAMRDVYIVQDLMETDLYKLLK	134
Db	61	PFEHQYTCQRTLEIQILLRFRHENVIGIRDILRAPLEAMRDVYIVQDLMETDLYKLLK	120
QY	135	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	194
Db	121	SQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDLKI CDFGLARIADPE	180
QY	195	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRRPIFPKGHYLDOL	254
Db	181	HDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDWSVGCILAEMLSNRRPIFPKGHYLDOL	240
QY	255	NHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVAAWAKLFPKSDSKALDLDRLMTFN	314
Db	241	NHILGILGSPSQEDLNCIINMKARNYLSQPSKTKVAAWAKLFPKSDSKALDLDRLMTFN	300
QY	315	PNKRIT-----VAEPPTFAMELDLDPKRLKELIFQETARFQPG	354
Db	301	PNKRITVEEALAHPLYEQYYDPTDEPVAEPPFTFDMELDDL PKRLKELIFQETARFQPG	360
QY	355	VLEAP 359	
Db	361	APEAP 365	

Search completed: February 6, 2006, 15:36:49
Job time : 37.2104 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:35:46 ; Search time 125.184 Seconds
(without alignments)
1198.238 Million cell updates/sec

Title: US-10-623-108-4

Perfect score: 1879

Sequence: 1 MAAAAAGGGGGPRTEGV.....LKELIFQETARFQGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications_AA_Main:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pgp:*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pgp:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pgp:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pgp:*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pgp:*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pgp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1879	100.0	359	5	US-10-623-108-4
2	1879	100.0	359	5	US-10-623-108-8
3	1873	99.7	359	4	US-10-114-270-110
4	1859	98.9	379	4	US-10-233-448-7
5	1859	98.9	379	4	US-10-174-794-14
6	1859	98.9	379	4	US-10-394-322A-19
7	1859	98.9	379	4	US-10-701-490-8
8	1859	98.9	379	5	US-10-735-118-3
9	1859	98.9	379	5	US-10-623-108-6
10	1859	98.9	379	5	US-10-945-684-2
11	1853	98.6	379	4	US-10-060-065-16
12	1853	98.6	379	4	US-10-059-585-37
13	1853	98.6	379	4	US-10-170-663-25
14	1853	98.6	379	5	US-10-893-072-25
15	1853	98.6	631	4	US-10-072-036-39
16	1787	95.1	624	4	US-10-072-036-57
17	1734	92.3	355	4	US-10-461-402-8
18	1599	85.1	335	5	US-10-623-108-2
19	1550.5	82.5	392	5	US-10-732-923-1420
20	1548.5	82.4	380	4	US-10-461-402-9
21	1538.5	81.9	360	4	US-10-171-311-127
22	1538.5	81.9	360	4	US-10-233-448-8
23	1538.5	81.9	360	4	US-10-301-822-115
24	1538.5	81.9	360	4	US-10-394-322A-20
25	1538.5	81.9	360	5	US-10-941-635-56
26	1532.5	81.6	360	4	US-10-664-421-56
27	1528.5	81.3	358	5	US-10-945-684-1

28	1528.5	81.3	358	5	US-10-451-291B-10	Sequence 10, Appl
29	1528.5	81.3	604	4	US-10-072-036-59	Sequence 59, Appl
30	1528.5	81.3	605	4	US-10-072-036-41	Sequence 41, Appl
31	1523	81.1	356	4	US-10-664-421-19	Sequence 19, Appl
32	1520	80.9	355	5	US-10-941-635-19	Sequence 19, Appl
33	1518	80.8	348	5	US-10-756-149-4856	Sequence 4856, Ap
34	1509.5	80.3	360	4	US-10-170-663-2	Sequence 2, Appl
35	1509.5	80.3	360	5	US-10-893-072-2	Sequence 2, Appl
36	1508	80.3	369	5	US-10-732-923-1421	Sequence 1421, Ap
37	1445	76.9	351	5	US-10-732-923-1406	Sequence 1406, Ap
38	1418	75.5	369	5	US-10-732-923-1456	Sequence 1456, Ap
39	1384.5	73.7	395	5	US-10-732-923-1424	Sequence 1424, Ap
40	1371.5	73.0	362	5	US-10-732-923-1418	Sequence 1418, Ap
41	1317	70.1	376	4	US-10-369-493-5689	Sequence 5689, Ap
42	944	50.2	352	5	US-10-732-923-1357	Sequence 1357, Ap
43	944	50.2	352	5	US-10-732-923-1367	Sequence 1367, Ap
44	941	50.1	355	5	US-10-732-923-1495	Sequence 1495, Ap
45	940	50.0	353	5	US-10-732-923-1492	Sequence 1492, Ap

ALIGNMENTS

RESULT 1

US-10-623-108-4
; Sequence 4, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-108-4

Query Match 100.0%; Score 1879; DB 5; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.3e-147;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAAAAGGGGGPRTEGVGPGVGEVENVKQPPDVGPRTYQLOVIGEGAYGMVSSAY	60
Db	1	MAAAAAGGGGGPRTEGVGPGVGEVENVKQPPDVGPRTYQLOVIGEGAYGMVSSAY	60
Qy	61	DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLASTLEAMRDVYI	120
Db	61	DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLASTLEAMRDVYI	120
Qy	121	VQDLMTDYLKLLKSQOLSNNDHICYFLYQILRLGKVIHNSANVLRDLKPSNLINTTCDL	180
Db	121	VQDLMTDYLKLLKSQOLSNNDHICYFLYQILRLGKVIHNSANVLRDLKPSNLINTTCDL	180
Qy	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDISWVGCI LAEMLS	240
Db	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDISWVGCI LAEMLS	240
Qy	241	NRPIFFGKYHLDQNLHILGILGSPSQEDLNCIIMKARNYLQSLPKTKVWAKLPKPSD	300
Db	241	NRPIFFGKYHLDQNLHILGILGSPSQEDLNCIIMKARNYLQSLPKTKVWAKLPKPSD	300
Qy	301	SKALDILDRMLTFNPNKRITVAEPEPTFAMELDDLPKRLKELI FORTARFQGVLEAP	359
Db	301	SKALDILDRMLTFNPNKRITVAEPEPTFAMELDDLPKRLKELI FORTARFQGVLEAP	359

RESULT 2

US-10-623-108-8
; Sequence 8, Application US/10623108

```
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAL, KEN-SHOW
; TITLE OF INVENTION: HUMAN SMAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-108-8

Query Match      100.0%; Score 1879; DB 5; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.3e-147;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAAQGGGGGPRRTGEGVGPGEVEMVKGPDPVGPRTQLOYGEGAYGMVSSAY 60
   |||
Db 1 MAAAAQGGGGGPRRTGEGVGPGEVEMVKGPDPVGPRTQLOYGEGAYGMVSSAY 60
   |||

QY 61 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||
Db 61 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||

QY 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKGYIHSANVLRDLKPSNLLINTTCDL 180
   |||
Db 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKGYIHSANVLRDLKPSNLLINTTCDL 180
   |||

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
   |||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
   |||

QY 241 NRPIFGKHYLDQNLNHLIGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||
Db 241 NRPIFGKHYLDQNLNHLIGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||

QY 301 SKALDILLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||
Db 301 SKALDILLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||

RESULT 3
US-10-114-270-110
; Sequence 110, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Zhaozhong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
```

```
; APPLICANT: Liete, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: Macdougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 110
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-110

Query Match      99.7%; Score 1873; DB 4; Length 359;
Best Local Similarity 99.7%; Pred. No. 1.4e-146;
Matches 358; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAAAQGGGGGPRRTGEGVGPGEVEMVKGPDPVGPRTQLOYGEGAYGMVSSAY 60
   |||
Db 1 MAAAAQGGGGGPRRTGEGVGPGEVEMVKGPDPVGPRTQLOYGEGAYGMVSSAY 60
   |||

QY 61 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||
Db 61 DHVTRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDILRASTLEAMRDVYI 120
   |||

QY 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKGYIHSANVLRDLKPSNLLINTTCDL 180
   |||
Db 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLKGYIHSANVLRDLKPSNLLINTTCDL 180
   |||

QY 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
   |||
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSDIWSVGCILAEMLS 240
   |||

QY 241 NRPIFGKHYLDQNLNHLIGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||
Db 241 NRPIFGKHYLDQNLNHLIGLSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLPKSD 300
   |||

QY 301 SKALDILLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||
Db 301 SKALDILLDRMLTFNPNKRITVAEPTTFAMELDDLKPERLKELIFOETARFQGVLEAP 359
   |||

RESULT 4
US-10-233-448-7
; Sequence 7, Application US/10233448
; Publication No. US20030109419A1
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
```

Matches	359;	Conservative	0;	Mismatches	0;	Indels	20;	Gaps	1;
Qy	1	MAAAAAQGGGCGGPRRTGCGVGPVCGEVEVMVKQGPFDVCGPRYTQLOQYIGEGAYGMVSSAY	60						
Db	1	MAAAAAQGGGCGGPRRTGCGVGPVCGEVEVMVKQGPFDVCGPRYTQLOQYIGEGAYGMVSSAY	60						
Qy	61	DHVRKTRVAIKKISPEHQTYCQRTLRREIQILRRFHEHNVIGIRDILRASTLEAMRDVYI	120						
Db	61	DHVRKTRVAIKKISPEHQTYCQRTLRREIQILRRFHEHNVIGIRDILRASTLEAMRDVYI	120						
Qy	121	VDQMETDLYKLLKSQQLSNDHI CYFLYQIILRGLKVIHSA NVLHRDLKPSNLLINTTCDL	180						
Db	121	VDQMETDLYKLLKSQQLSNDHI CYFLYQIILRGLKVIHSA NVLHRDLKPSNLLINTTCDL	180						
Qy	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKG YTKSIDIWSVGCILAEMLS	240						
Db	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKG YTKSIDIWSVGCILAEMLS	240						
Qy	241	NRPIPFKGHYLDQNLHILGILSPSQEDLNCIINMKARNYQSLPSKTKVAMAKLFPKSD	300						
Db	241	NRPIPFKGHYLDQNLHILGILSPSQEDLNCIINMKARNYQSLPSKTKVAMAKLFPKSD	300						
Qy	301	SKALDLLDRMLTFNPNKRIIT-----VAEPPTFAMELDDLPKRL	340						
Db	301	SKALDLLDRMLTFNPNKRIITVEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKRL	360						
Qy	341	KELIQETARFQGVLEAP	359						
Db	361	KELIQETARFQGVLEAP	379						

RESULT 6

US-10-394-322A-19

Sequence 19, Application US/10394322A

Publication No. US20030232391A1

GENERAL INFORMATION:

APPLICANT: SINESIS PHARMACEUTICALS, INC.

APPLICANT: Prescott, John C.

TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS

FILE REFERENCE: 39750-0006 US

CURRENT APPLICATION NUMBER: US/10/394,322A

CURRENT FILING DATE: 2003-03-20

PRIOR APPLICATION NUMBER: US 60/366,892

PRIOR FILING DATE: 2002-03-21

NUMBER OF SEQ ID NOS: 70

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 19

LENGTH: 379

TYPE: PPT

ORGANISM: Homo sapiens

US-10-394-322A-19

Query Match 98.9%; Score 1859; DB 4; Length 379;

Best Local Similarity 94.7%; Pred. No. 2.1e-145;

Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

Qy	1	MAAAAAQGGGCGGPRRTGCGVGPVCGEVEVMVKQGPFDVCGPRYTQLOQYIGEGAYGMVSSAY	60						
Db	1	MAAAAAQGGGCGGPRRTGCGVGPVCGEVEVMVKQGPFDVCGPRYTQLOQYIGEGAYGMVSSAY	60						
Qy	61	DHVRKTRVAIKKISPEHQTYCQRTLRREIQILRRFHEHNVIGIRDILRASTLEAMRDVYI	120						
Db	61	DHVRKTRVAIKKISPEHQTYCQRTLRREIQILRRFHEHNVIGIRDILRASTLEAMRDVYI	120						
Qy	121	VDQMETDLYKLLKSQQLSNDHI CYFLYQIILRGLKVIHSA NVLHRDLKPSNLLINTTCDL	180						
Db	121	VDQMETDLYKLLKSQQLSNDHI CYFLYQIILRGLKVIHSA NVLHRDLKPSNLLINTTCDL	180						
Qy	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKG YTKSIDIWSVGCILAEMLS	240						
Db	181	KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLNSKG YTKSIDIWSVGCILAEMLS	240						
Qy	241	NRPIPFKGHYLDQNLHILGILSPSQEDLNCIINMKARNYQSLPSKTKVAMAKLFPKSD	300						

```
Db 241 NRPIFGKHYLDQLNHLGILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 7
US-10-701-490-8
; Sequence 8, Application US/10701490
; Publication No. US20040106141A1
; GENERAL INFORMATION:
; APPLICANT: PAUL S. MISCHER
; APPLICANT: CHARLES L. SAWYERS
; APPLICANT: BRADLEY L. SMITH
; APPLICANT: KATHERINE CROSBY
; TITLE OF INVENTION: METHODS AND MATERIALS FOR EXAMINING
; TITLE OF INVENTION: PATHWAYS ASSOCIATED WITH GLIOBLASTOMA PROGRESSION
; FILE REFERENCE: G&C 30435.148USU1
; CURRENT APPLICATION NUMBER: US/10/701,490
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/423,777
; PRIOR FILING DATE: 2002-11-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-701-490-8

Query Match 98.9%; Score 1859; DB 4; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

Qy 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCIIAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCIIAEMLS 240
Qy 241 NRPIFGKHYLDQLNHLGILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSD 300
Db 241 NRPIFGKHYLDQLNHLGILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 8
US-10-735-118-3
; Sequence 3, Application US/10735118
; Publication No. US20040248151A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Bacue, Sarah S.
; APPLICANT: Smith, Bradley L.
; TITLE OF INVENTION: METHOD FOR PREDICTING THE RESPONSE TO HER2-DIRECTED THERAPY
; FILE REFERENCE: 6270-701.201
; CURRENT APPLICATION NUMBER: US/10/735,118
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 60/370,473
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/432,943
; PRIOR FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 10/408,520
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-735-118-3

Query Match 98.9%; Score 1859; DB 5; Length 379;
Best Local Similarity 94.7%; Pred. No. 2.1e-145;
Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

Qy 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGEPRTTEGVGPGVEVMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQTYCQRTLREIQILLRFRHENVIGIRDLRASTLEAMRDVYI 120
Qy 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCIIAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCIIAEMLS 240
Qy 241 NRPIFGKHYLDQLNHLGILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSD 300
Db 241 NRPIFGKHYLDQLNHLGILGSPSQEDLNCIINMKARNYLSLPSKTKVAVAKLFPKSD 300
Qy 301 SKALDLDLRLMTFNPKNRIT-----VAEPPTFAMELDDLPKERL 340
Db 301 SKALDLDLRLMTFNPKNRITVEEALAHPLYEQYDPTDEPVAEPPTFAMELDDLPKERL 360
Qy 341 KELIFQETARFQGVLEAP 359
Db 361 KELIFQETARFQGVLEAP 379

RESULT 9
US-10-623-108-6
; Sequence 6, Application US/10623108
; Publication No. US20050013817A1
; GENERAL INFORMATION:
; APPLICANT: DAI, KEN-SHOW
; TITLE OF INVENTION: HUMAN SNAPK3-RELATED GENE VARIANTS ASSOCIATED WITH CANCERS
; FILE REFERENCE: U 014726-8
; CURRENT APPLICATION NUMBER: US/10/623,108
; CURRENT FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-108-6

Query Match 98.9%; Score 1859; DB 5; Length 379;
```

Best Local Similarity 94.7%; Pred. No. 2.1e-145; Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;									
Qy	1	MAAAAQGGGGPRRTGEGVPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60						
Db	1	MAAAAQGGGGPRRTGEGVPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60						
Qy	61	DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI	120						
Db	61	DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI	120						
Qy	121	VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL	180						
Db	121	VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL	180						
Qy	181	KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240						
Db	181	KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240						
Qy	241	NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSD	300						
Db	241	NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSD	300						
Qy	301	SKALDLLDRMLTFNPNKRIT-----VAEPTTFAMELDDLPKRL	340						
Db	301	SKALDLLDRMLTFNPNKRITVEEALAHPLYEQYDPTDFVAEPTTFAMELDDLPKRL	360						
Qy	341	KELIFQETARFQGVLEAP	359						
Db	361	KELIFQETARFQGVLEAP	379						
RESULT 10									
US-10-945-684-2									
; Sequence 2, Application US/10945684									
; Publication No. US20050095657A1									
; GENERAL INFORMATION:									
; APPLICANT: Emory University									
; TITLE OF INVENTION: Methods and Kits for Detecting Proteins									
; FILE REFERENCE: 50508-2280									
; CURRENT APPLICATION NUMBER: US/10/945,684									
; CURRENT FILING DATE: 2004-09-21									
; NUMBER OF SEQ ID NOS: 8									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 2									
; TYPE: PRT									
; ORGANISM: Homo Sapiens									
US-10-945-684-2									
Query Match 98.9%; Score 1859; DB 5; Length 379;									
Best Local Similarity 94.7%; Pred. No. 2.1e-145; Matches 359; Conservative 0; Mismatches 0; Indels 20; Gaps 1;									
Qy	1	MAAAAQGGGGPRRTGEGVPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60						
Db	1	MAAAAQGGGGPRRTGEGVPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60						
Qy	61	DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI	120						
Db	61	DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI	120						
Qy	121	VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL	180						
Db	121	VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL	180						
Qy	181	KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240						
Db	181	KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240						
Qy	241	NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSD	300						
Db	241	NRPIFPKGHYLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLFPKSD	300						
US-10-060-065-16									
; Sequence 16, Application US/10060065									
; Publication No. US20030017480A1									
; GENERAL INFORMATION:									
; APPLICANT: Toshio Ota									
; APPLICANT: Takao Isogai									
; APPLICANT: Tetsuo Nishikawa									
; APPLICANT: Koji Hayashi									
; APPLICANT: Kaoru Otsuka									
; APPLICANT: Jun-ichi Yamamoto									
; APPLICANT: Shizuko Ishii									
; APPLICANT: Tomoyasu Sugiyama									
; APPLICANT: Ai Wakamatsu									
; APPLICANT: Keiichi Nagai									
; APPLICANT: Tetsuji Otsuki									
; APPLICANT: Shin-ichi Funahashi									
; APPLICANT: Chiaki Senoo									
; APPLICANT: Jun-ichi Nezu									
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE									
; FILE REFERENCE: 06501-099002									
; CURRENT APPLICATION NUMBER: US/10/060,065									
; CURRENT FILING DATE: 2002-01-29									
; PRIOR APPLICATION NUMBER: PCT/JP00/05061									
; PRIOR FILING DATE: 2000-07-28									
; PRIOR APPLICATION NUMBER: US 60/159,590									
; PRIOR FILING DATE: 1999-10-18									
; PRIOR APPLICATION NUMBER: US 60/183,322									
; PRIOR FILING DATE: 2000-02-17									
; PRIOR APPLICATION NUMBER: JP 11-248036									
; PRIOR FILING DATE: 1999-07-29									
; PRIOR APPLICATION NUMBER: JP 2000-118776									
; PRIOR FILING DATE: 2000-01-11									
; PRIOR APPLICATION NUMBER: JP 2000-183767									
; PRIOR FILING DATE: 2000-05-02									
; PRIOR APPLICATION NUMBER: JP 2000-241899									
; PRIOR FILING DATE: 2000-06-09									
; NUMBER OF SEQ ID NOS: 43									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 16									
; LENGTH: 379									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-060-065-16									
Query Match 98.6%; Score 1853; DB 4; Length 379;									
Best Local Similarity 94.5%; Pred. No. 6.6e-145; Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;									
Qy	1	MAAAAQGGGGPRRTGEGVPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60						
Db	1	MAAAAQGGGGPRRTGEGVPGVGEVEMVKQPPDVGPRYTQLOVIGEGAYGMVSSAY	60						
Qy	61	DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI	120						
Db	61	DHVKTRVAIKKISPEFHQTYCQRTLREIQILLRFHENVIGIRDIRASTLEAMRDVYI	120						
Qy	121	VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL	180						
Db	121	VQDLMETDLYKLLKSQOLSNDHICYFLYQILRLGKYIHSANVLRDLKPSNLLINTTCDL	180						
Qy	181	KICDFGLARIADPEHDTGFLTEYVATRWYRAPEIMLSKGYTKSIDWSVGCILAEMLS	240						

Db 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240

Qy 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSD 300

Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSD 300

Qy 301 SKALDLDRLMTFNPKNKRIT-----VAEPPFTFAMELDDLPKRL 340

Db 301 SKALDLDRLMTFNPKNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKRL 360

Qy 341 KELIFOETARFQGVLEAP 359

Db 361 KELIFOETARFQGVLEAP 379

RESULT 12

US-10-059-585-37

Sequence 37, Application US/10059585

Publication No. US20030082776A1

GENERAL INFORMATION:

APPLICANT: Ota, Toshio

APPLICANT: Isogai, Takao

APPLICANT: Nishikawa, Tetsuo

APPLICANT: Hayashi, Koji

APPLICANT: Otsuka, Kaoru

APPLICANT: Yamamoto, Jun-ichi

APPLICANT: Ishii, Shizuko

APPLICANT: Sugiyama, Tomoyasu

APPLICANT: Wakamatsu, Ai

APPLICANT: Nagai, Keiichi

APPLICANT: Otsuki, Tetsuji

APPLICANT: Funahashi, Shin-Ichi

APPLICANT: Senoo, Chiaki

APPLICANT: Nezu, Jun-ichi

TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN

TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE

FILE REFERENCE: 06501-098001

CURRENT APPLICATION NUMBER: US/10/059,585

CURRENT FILING DATE: 2002-01-29

PRIOR APPLICATION NUMBER: PCT/JP00/05060

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: US 60/183,322

PRIOR FILING DATE: 2000-02-17

PRIOR APPLICATION NUMBER: US 60/159,590

PRIOR FILING DATE: 1999-10-18

PRIOR APPLICATION NUMBER: JP 2000-118776

PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: JP 2000-183767

PRIOR FILING DATE: 2000-05-02

PRIOR APPLICATION NUMBER: JP 11-248036

PRIOR FILING DATE: 1999-07-29

NUMBER OF SEQ ID NOS: 64

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 37

LENGTH: 379

TYPE: PRT

ORGANISM: Homo sapiens

US-10-059-585-37

Query Match 98.6%; Score 1853; DB 4; Length 379;

Best Local Similarity 94.5%; Pred. No. 6.6e-145;

Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

Qy 1 MAAAAQGGGGGPRRTGEGVGPVGEVEMVKGPFDVGPRTYQIYIGEGAYGMVSSAY 60

Db 1 MAAAAQGGGGGPRRTGEGVGPVGEVEMVKGPFDVGPRTYQIYIGEGAYGMVSSAY 60

Qy 61 DHVKTRVAKKISPEHYQTCORTREIQILLRFRHENVIGRDLIRASTLEAMRDVYI 120

Db 61 DHVKTRVAKKISPEHYQTCORTREIQILLRFRHENVIGRDLIRASTLEAMRDVYI 120

Qy 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLNTTCDL 180

Db 121 VQDLMETDLYKLLKSQQLSNDHICYFLYQILRLGLKYIHSANVLHRLDKPSNLLNTTCDL 180

Qy 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240

Db 181 KICDGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS 240

Qy 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSD 300

Db 241 NRPIFGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLPSKTKVAVAKLFPKSD 300

Qy 301 SKALDLDRLMTFNPKNKRIT-----VAEPPFTFAMELDDLPKRL 340

Db 301 SKALDLDRLMTFNPKNKRITVEEALAHPLYEQYDPTDEPVAEPPFTFAMELDDLPKRL 360

Qy 341 KELIFOETARFQGVLEAP 359

Db 361 KELIFOETARFQGVLEAP 379

RESULT 13

US-10-170-663-25

Sequence 25, Application US/10170663

Publication No. US20030165899A1

GENERAL INFORMATION:

APPLICANT: Su, Michael Shin-San

Fox, Ted

Wilson, Keith Phillip

Germann, Ursula A.

TITLE OF INVENTION: Methods For Designing Inhibitors of Serine/Threonine Kinases and Tyrosine Kinase

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Neave

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/170,663

FILING DATE: 12-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/025,580

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Haley, James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: VPI 97-104

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-10-170-663-25

Query Match 98.6%; Score 1853; DB 4; Length 379;

Best Local Similarity 94.5%; Pred. No. 6.6e-145;

Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

Qy 1 MAAAAQGGGGGPRRTGEGVGPVGEVEMVKGPFDVGPRTYQIYIGEGAYGMVSSAY 60

Db 1 MAAAAAQQGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Qy 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 240
Qy 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLPKSD 300
Db 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLPKSD 300
Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPTFFAMELDDLPKRL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDBEVAEPTFFAMELDDLPKRL 360
Qy 341 KELIFQETARFQPGVLEAP 359
Db 361 KELIFQETARFQPGVLEAP 379

RESULT 14
US-10-893-072-25
; Sequence 25, Application US/10893072
; Publication No. US20040259166A1
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; Fox, Ted
; Wilson, Keith Phillip
; Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/893,072
; FILING DATE: 15-Jul-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/642,749
; FILING DATE: 18-Aug-2000
; APPLICATION NUMBER: US/09/025,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-893-072-25
Query Match 98.6%; Score 1853; DB 5; Length 379;
Best Local Similarity 94.5%; Pred. No. 6.6e-145; Indels 20; Gaps 1;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;
Qy 1 MAAAAAQQGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Db 1 MAAAAAQQGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Qy 61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Db 61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFRHENVIGIRDLASTLEAMRDVYI 120
Qy 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Db 121 VQDLMTDLYKLLKSQQLSNDHICYFLYQILRGLKYIHSANVLRDLKPSNLLINTTCDL 180
Qy 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 240
Db 181 KICDFGLARIADPEHDHTGFLTEYVATRWYRAPEIMLSNGYTKSIDINSVGCILAEMLS 240
Qy 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLPKSD 300
Db 241 NRPIFGKHVLDQNLHILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAAWAKLPKSD 300
Qy 301 SKALDLDRLMTFNPKNRIT-----VAEPTFFAMELDDLPKRL 340
Db 301 SKALDLDRLMTFNPKNRITVEEALAHPLYEQYYDPTDBEVAEPTFFAMELDDLPKRL 360
Qy 341 KELIFQETARFQPGVLEAP 359
Db 361 KELIFQETARFQPGVLEAP 379

RESULT 15
US-10-072-036-39
; Sequence 39, Application US/10072036
; Publication No. US20030082564A1
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP
; APPLICANT: Sara BJRON
; APPLICANT: Soren TULLIN
; APPLICANT: Kasper ALMHOLT
; APPLICANT: Kurt SCUDDER
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; TITLE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0120P
; CURRENT APPLICATION NUMBER: US/10/072,036
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/417,197
; PRIOR FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 39
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-Erk1 fusion construct
US-10-072-036-39

Query Match 98.6%; Score 1853; DB 4; Length 631;
Best Local Similarity 94.5%; Pred. No. 1.2e-144;
Matches 358; Conservative 0; Mismatches 1; Indels 20; Gaps 1;
Qy 1 MAAAAAQQGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 60
Db 253 MAAAAAQQGGGGPRRTEGVGPGVGEVEMVKGPFDVGPRTYQLOVIGEGAYGMVSSAY 312
Qy 61 DHVRKTRVAIKKISPEHQYTCQRTLEIQILLRFRHENVIGIRDLASTLEAMRDVYI 120

313	Db	DHVRKTRVAIKKISPEHQTYCQRTREIQILLRFRHENVIGIRDILRASTLEARNRDVYI	372
121	Qy	VQDLMETDLYKLLKSQQLSNDHIHCVFLYQIILRGLKYIHSANVLHRDLKPSNLLINTTCDL	180
373	Db	VQDLMETDLYKLLKSQQLSNDHIHCVFLYQIILRGLKYIHSANVLHRDLKPSNLLSNTTCDL	432
181	Qy	KICDFGLARIADPEHDHTGFTTEYVATRWYRAPEIMLNSKGYTKSIDISWVGCIILAEMLS	240
433	Db	KICDFGLARIADPEHDHTGFTTEYVATRWYRAPEIMLNSKGYTKSIDISWVGCIILAEMLS	492
241	Qy	NRPIFPGKHYLDQLNHIILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAWAKLPKPSD	300
493	Db	NRPIFPGKHYLDQLNHIILGILGSPSQEDLNCIINMKARNYLQSLPSKTKVAWAKLPKPSD	552
301	Qy	SKALDLDLRLMTFENPKRIT-----VAEEFFTFAMELDDLDLPKERL	340
553	Db	SKALDLDLRLMTFENPKRITVEEALAHPLYEQYDPTDEPVAEEFFTFAMELDDLDLPKERL	612
341	Qy	KELIFQETARFQPGVLEAP	359
613	Db	KELIFQETARFQPGVLEAP	631

Search completed: February 6, 2006, 15:40:57
Job time : 126.184 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2006, 15:37:06 ; Search time 10.8631 Seconds
(without alignments)
387.254 Million cell updates/sec

Title: US-10-623-108-4

Perfect score: 1879

Sequence: 1 MAAAAAGGGGGPRRTGV.....LKELIFQETARFGVLEAP 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA_New:
- 1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pap:*
 - 2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pap:*
 - 3: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pap:*
 - 4: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pap:*
 - 5: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pap:*
 - 6: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pap:*
 - 7: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pap:*
 - 8: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1853	98.6	379	7	US-11-109-156-16
2	1538.5	81.9	360	6	US-10-878-556A-134
3	1538.5	81.9	360	7	US-11-186-284-115
4	758.5	40.4	360	6	US-10-886-329-1
5	704.5	37.5	365	6	US-10-770-726-69
6	629	33.5	422	6	US-10-857-780-21
7	629	33.5	422	7	US-11-127-817-18
8	629	33.5	426	7	US-11-127-817-20
9	629	33.5	464	7	US-11-127-817-19
10	491	26.1	460	7	US-11-024-959-384
11	490.5	26.1	294	7	US-11-024-959-263
12	481.5	25.6	294	7	US-11-024-959-265
13	479.5	25.5	304	7	US-11-024-959-264
14	477.5	25.4	302	7	US-11-024-959-262
15	472.5	25.1	520	7	US-11-024-959-272
16	471.5	25.1	298	6	US-10-770-726-51
17	458	24.4	277	7	US-11-127-817-21
18	451.5	24.0	346	6	US-10-770-726-55
19	450.5	24.0	297	7	US-10-770-726-48
20	450.5	24.0	297	7	US-11-109-156-11
21	449.5	23.9	795	6	US-10-770-726-49
22	449	23.9	496	6	US-10-770-726-72
23	448.5	23.9	555	7	US-11-024-959-270
24	445.5	23.7	509	7	US-11-024-959-393
25	443.5	23.6	463	7	US-11-024-959-395

26	442	23.5	292	6	US-10-770-726-53	Sequence 53, Appl
27	441	23.5	483	6	US-10-451-375-12	Sequence 12, Appl
28	439	23.4	476	7	US-11-024-959-385	Sequence 385, App
29	438	23.3	433	6	US-10-770-726-65	Sequence 65, Appl
30	437	23.3	313	7	US-11-024-959-267	Sequence 267, App
31	433	23.0	706	7	US-11-024-959-269	Sequence 269, App
32	431	22.9	477	7	US-11-024-959-261	Sequence 261, App
33	426.5	22.7	330	6	US-10-786-065-8	Sequence 8, Appl
34	424.5	22.6	330	6	US-10-786-065-10	Sequence 10, Appl
35	423.5	22.5	330	6	US-10-786-065-10	Sequence 268, App
36	421	22.4	583	7	US-11-024-959-268	Sequence 5, Appl
37	419	22.3	435	6	US-10-786-065-5	Sequence 273, App
38	419	22.3	718	7	US-11-024-959-273	Sequence 386, App
39	415	22.1	302	7	US-11-024-959-386	Sequence 391, App
40	414	22.0	845	7	US-11-024-959-391	Sequence 271, App
41	412	21.9	469	7	US-11-024-959-271	Sequence 392, App
42	405	21.6	797	7	US-11-024-959-392	Sequence 52, Appl
43	401.5	21.4	303	6	US-10-770-726-52	Sequence 266, App
44	398.5	21.2	305	7	US-11-024-959-266	Sequence 387, App
45	395	21.0	302	7	US-11-024-959-387	

ALIGNMENTS

RESULT 1
US-11-109-156-16
; Sequence 16, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT FILING DATE: 2005-04-19
; PRIOR FILING DATE: 2002-01-29
; PRIOR FILING DATE: 2002-01-29
; PRIOR FILING DATE: 2000-07-28
; PRIOR FILING DATE: 1999-10-18
; PRIOR FILING DATE: 2000-02-17
; PRIOR FILING DATE: 2000-02-17
; PRIOR FILING DATE: 1999-07-29
; PRIOR FILING DATE: 2000-01-11
; PRIOR FILING DATE: 2000-01-11
; PRIOR FILING DATE: 2000-05-02
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-16

Query Match		98.6%;	Score 1853;	DB 7;	Length 379;
Best Local Similarity		94.5%;	Pred. No. 4.4e-157;		
Matches		358;	Conservative	0;	Mismatches 1; Indels 20; Gaps 1;
QY	1	MAAAAQGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAQGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
QY	61	DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRSTLEAMRDVYI	120		
DB	61	DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRSTLEAMRDVYI	120		
QY	121	VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	180		
DB	121	VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	180		
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
DB	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
QY	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	300		
DB	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	300		
QY	301	SKALDLLDRMLTFNPNKRIT-----VAEPTTFAMELDDLPKERL	340		
DB	301	SKALDLLDRMLTFNPNKRITVEBALAHPLYEQYDPTDFVAEPTTFAMELDDLPKERL	360		
QY	341	KELIFQETARFQGVLEAP	359		
DB	361	KELIFQETARFQGVLEAP	379		
RESULT 2					
US-10-878-556A-134					
; Sequence 134, Application US/10878556A					
; Publication No. US20050266399A1					
; GENERAL INFORMATION:					
; APPLICANT: Hoffmann La-Roche Inc.					
; TITLE OF INVENTION: HCV regulated protein expression					
; FILE REFERENCE: 21762					
; CURRENT APPLICATION NUMBER: US/10/878, 556A					
; CURRENT FILING DATE: 2004-06-28					
; NUMBER OF SEQ ID NOS: 199					
; SOFTWARE: PatentIn version 3.1					
; SEQ ID NO 134					
; LENGTH: 360					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
; PUBLICATION INFORMATION:					
; DATABASE ACCESSION NUMBER: sw_hum/mk01_human					
; DATABASE ENTRY DATE: 1992-12-01					
US-10-878-556A-134					
Query Match		81.9%;	Score 1538.5;	DB 6;	Length 360;
Best Local Similarity		79.1%;	Pred. No. 3.3e-129;		
Matches		296;	Conservative	18;	Mismatches 23; Indels 37; Gaps 2;
QY	1	MAAAAQGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAAGAG-----PEMVRQGVDPGPRYTNLISYIGEGAYGMVCSAY	43		
QY	61	DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRSTLEAMRDVYI	120		
DB	44	DNVKNRVAIKKISPEHQTYCORTLREIKILLRFRHENIIGINDIIRAPTIQMKDXYI	103		
QY	121	VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	180		
DB	104	VQDLMETDLYKLLKQHLNSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	163		
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
DB	164	KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	223		
QY	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	300		
DB	224	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	293		

Query Match		81.9%;	Score 1538.5;	DB 7;	Length 360;
Best Local Similarity		79.1%;	Pred. No. 3.3e-129;		
Matches		296;	Conservative	18;	Mismatches 23; Indels 37; Gaps 2;
QY	1	MAAAAQGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAAGAG-----PEMVRQGVDPGPRYTNLISYIGEGAYGMVCSAY	43		
QY	61	DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRSTLEAMRDVYI	120		
DB	44	DNVKNRVAIKKISPEHQTYCORTLREIKILLRFRHENIIGINDIIRAPTIQMKDXYI	103		
QY	121	VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	180		
DB	104	VQDLMETDLYKLLKQHLNSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	163		
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
DB	164	KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	223		
QY	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	300		
DB	224	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	293		
RESULT 3					
US-11-186-284-115					
; Sequence 115, Application US/11186284					
; Publication No. US20050266493A1					
; GENERAL INFORMATION:					
; APPLICANT: Millennium Pharmaceuticals, Inc.					
; APPLICANT: Berger, Allison					
; APPLICANT: Guillemette, Tracy L.					
; APPLICANT: Kamatkar, Shubhangi					
; APPLICANT: Schlegel, Robert					
; APPLICANT: Monahan, John E.					
; APPLICANT: Thibodeau, Stephen N.					
; APPLICANT: Burgart, Lawrence J.					
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND					
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND					
; FILE REFERENCE: MPM01-029P2RNM					
; CURRENT APPLICATION NUMBER: US/11/186,284					
; CURRENT FILING DATE: 2005-07-21					
; PRIOR APPLICATION NUMBER: US/10/301,822					
; PRIOR FILING DATE: 2002-11-21					
; PRIOR APPLICATION NUMBER: US 60/339,971					
; PRIOR FILING DATE: 2001-12-10					
; PRIOR APPLICATION NUMBER: US 60/361,978					
; PRIOR FILING DATE: 2002-03-05					
; PRIOR APPLICATION NUMBER: US 60/381,988					
; PRIOR FILING DATE: 2002-08-20					
; NUMBER OF SEQ ID NOS: 228					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 115					
; LENGTH: 360					
; TYPE: PRT					
; ORGANISM: Homo Sapiens					
US-11-186-284-115					
Query Match		81.9%;	Score 1538.5;	DB 7;	Length 360;
Best Local Similarity		79.1%;	Pred. No. 3.3e-129;		
Matches		296;	Conservative	18;	Mismatches 23; Indels 37; Gaps 2;
QY	1	MAAAAQGGGGPRRTGEGVGPGEVEMVKGPDPDVGPRYTQLOYIGEGAYGMVSSAY	60		
DB	1	MAAAAAGAG-----PEMVRQGVDPGPRYTNLISYIGEGAYGMVCSAY	43		
QY	61	DHVRKTRVAIKKISPEHQTYCORTLREIQILLRFRHENVIGIRDLRSTLEAMRDVYI	120		
DB	44	DNVKNRVAIKKISPEHQTYCORTLREIKILLRFRHENIIGINDIIRAPTIQMKDXYI	103		
QY	121	VQDLMETDLYKLLKSQOLSNNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	180		
DB	104	VQDLMETDLYKLLKQHLNSNDHICYFLYQILRLGKLYIHSANVLRDLKPSNLLINTTCDL	163		
QY	181	KICDFGLARIADPEHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	240		
DB	164	KICDFGLARVADPDHDHTGFLTTEYVATRWYRAPEIMLSKGYTKSIDIWSVGCILAEMLS	223		
QY	241	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	300		
DB	224	NRPIPGKHYLDQNLHILGILGSPSQEDLNCIINMKARNYLOSLSKTKVAVAKLPKSD	293		

Job time : 11.8631 secs

Db 63 IVALYDVWTE-----KGLTVFEFLDQDLKYLDAAGDNGLEPYTVKSYLQQLGIAP 117
Qy 157 IHSANVLHRLKPSNLLINTTCDLKICDFGLAR-IADPEHDHGTGFLTEYVATRWYRAPEI 215
Db 118 CHEHRLVHRLKPEQLNINMEGSLKGLADGLARAFGIPVRY-----THEVTLWYRAPDV 173
Qy 216 MLNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYDLQNLHILGILGSPSQEDLNCIINN 275
Db 174 LMSRKYSTQVDIWSVGCIFAEMVNGRPLPPGSSEQDQLLRIFKTLGTPSLKTWPGMAEL 233
Qy 276 KARYLQSLPSKYKVAWAKLFPKS-DSKALDLDRLMTFNPNGRIITVAE 323
Db 234 P--DFKDNFPKYVQSPFKICPKKLDKTKGLDLSRMLQYDPAKRISABQ 280

RESULT 15

US-11-024-959-272
; Sequence 272, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 272
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Eucalyptus sp.
US-11-024-959-272

Query Match 25.1%; Score 472.5; DB 7; Length 520;
Best Local Similarity 34.2%; Pred. No. 2e-34;
Matches 122; Conservative 56; Mismatches 116; Indels 63; Gaps 8;
Qy 42 YTLQYIGEGAYGMVSSAYDHVRKTRVAIKKI-SPPEHQTYCORTLREIQIILLRFRHNV 100
Db 25 FEKLEQIGEGTYGOVYMAKEKKTGEIVALKKIRMDNEREGFPITAIREIKLKLLHHNV 84
Qy 101 IGIRDILRASTLEAMRD-----VYIVDLMETDLYKLL--KSQQLSNDHICYFL 147
Db 85 IKLBEIVTSPGPEKDEGRPEGNKYGGIYVFEYMDHDLTGLADRFGRFSPVQIKCYM 144
Qy 148 YQILRGLKYIHSANVLHRLKPSNLLINTTCDLKICDFGLAR-IADPEHDHGTGFLTEYVAT 207
Db 145 RQLLTGLHYCHINQVLRHDIKGNSLLIDNEGNLKLADFGLAR--SPSNDHNANITNRVIT 202
Qy 208 RWTAPRIMLNSKGYTKSIDIWSVGCILAEMLSNRPIPPGKHLYDLQNLHILGILGSPSQE 267
Db 203 LWYRPPELLLGATKYGPAVDMWSVGCIFAELLHGKPIPPGKDEPEQLNKIFELCGAPDE- 261
Qy 268 DLNCIINWKARNYLQSLPSKYKVAW-----AKLFPKSPSKALDLDRLMTF 313
Db 262 -----INW-----PGVSKIPIYNNFKTRPMKRLREVFHRHLELLERMLTL 307
Qy 314 NPNKRITVAEBPFTFAMELDDLP-----KERLKELIPOETARFQ 352
Db 308 DPSORISAKDALDAEYFWADPLCPDKPSLPKYESSEHFEPTKKKQOQROHEETAKRQ 364

Search completed: February 6, 2006, 15:41:23

This Page Blank (uspto)